Eur J Forest Res (2012) 131:441–451
DOI 10.1007/s10342-011-0518-0

ORIGINAL PAPER

Estimating tree species diversity across geographic scales

Susanne Winter · Andreas Böck · Ronald E. McRoberts

Received: 22 July 2010 / Revised: 15 April 2011 / Accepted: 18 April 2011 / Published online: 12 May 2011

Abstract The relationship between number of species and area observed has been described using numerous approaches and has been discussed for more than a century. The general objectives of our study were fourfold: (1) to evaluate the behaviour of species–area curves across geographic scales, (2) to determine sample sizes necessary to produce acceptably precise estimates of tree species diversity, (3) to evaluate relationships among estimates of tree species diversity for local to large geographic scales, and (4) to determine whether the proportion of native tree species may be precisely estimated using sample sizes smaller than those necessary to estimate total tree species diversity. Such investigations are necessary to improve biodiversity monitoring at large scales. The analytical approaches included Monte Carlo bootstrap simulations and two geospatial methods and relied on a database populated with data for more than 12,000 national forest inventory plots (NFI) from 16 regional units, 13 European countries and three ecoprovinces of the United States of America (USA). Four primary results were obtained. First, tree species diversity may be precisely estimated using observations for a random subsample of 2,000–4,000 NFI plots. Second, large sample sizes are necessary to estimate tree species diversity for regional units or forest categories, except possibly for boreal forests for which the number of tree species is small. Third, estimates of the proportion of native tree species may be precisely estimated using tree species information for a random sample of approximately 30 NFI plots. Finally, our estimated species–area curves show that the curve shapes and relationships among estimates of tree species diversity at large scales clearly depend on the relative geographic location of the anchor regional unit (a European country or ecoprovince of the USA) relative to the other regional units. None of the well-known models for species–area curves adequately describes our results. The conclusion was that the uncertainties of the estimates reflect the unfavourable state of global biodiversity monitoring of species groups. The total numbers of tree species in our regional units, which cannot be adequately estimated, are small relative to other tree species-rich regions. Consequently, monitoring of tree species diversity is currently a highly uncertain enterprise at large scales.

Keywords Biogeography · Species–area curve · Forest category · Forest structure · International reporting · National forest inventory · Spatial autocorrelation

Abbreviations

NFI	National forest inventory
ln	Natural logarithmic to the base e with e approximately 2,71

Communicated by K. Puettmannn.

S. Winter (✉)
Department für Ökologie, Department for Ecology and Ecosystem Management, Geobotany, Technische Universität München, Hans-Carl-von-Carlowitz-Platz 2, 85354 Freising, Germany
e-mail: winter@wzw.tum.de

A. Böck
Technische Universität München, Fachgbiet Biostatistik, Zentrum für Mathematik, Am Forum 1, 85354 Freising, Germany
e-mail: andreas.boeck@tum.de

R. E. McRoberts
Northern Research Station, US Forest Service, Saint Paul, MN 55108, USA
e-mail: rmcroberts@fs.fed.us

Introduction

Common reporting on biodiversity is increasingly required to monitor biodiversity changes at large scales (Chirici et al. 2011; Boutin et al. 2009; Winter et al. 2008). This requirement is documented in international agreements, protocols and processes such as the Convention on Biological Diversity (CBD 1992), the Montréal Process (2009), and the Ministerial Conference on the Protection of Forests in Europe (e.g. MCPFE 1997, 2003). Detailed biodiversity monitoring is a prerequisite to managing and maintaining biodiversity (Brockway 1998), and knowledge of how biodiversity changes from local to large scales is necessary to interpret the results of harmonised biodiversity assessments (Chirici et al. 2011). Increasingly, biodiversity estimates are scaled up from local (alpha) and regional (gamma) to large-scale levels (Whittaker 1972, 1975, 1977; Whittaker et al. 2001). Local or alpha biodiversity is the diversity of very small habitats such as forest stands and is often studied using plots of sizes on the order of m^2 to hundreds of m^2 (Tomppo et al. 2010). Regional or gamma diversity focuses on total diversity at the landscape level and includes different habitats such as forest categories (EEA 2006), and large-scale diversity refers to geographic scales that include multiple landscapes or regions such as the entirety of Europe.

Schoener (1976) asserted that the relationship between increasing numbers of species and increasing area observed constitutes one of the few laws of ecology. This relationship dates back to Watson (1859) who constructed the first local- to large-scale species–area curve for plant species in Great Britain. Decades later, Arrhenius (1921) and Gleason (1922) described the relationship between the number of species and area observed using two theoretical modelling approaches: (1) the power model and (2) the exponential model. Both models have been discussed and refined by numerous researchers (Preston 1960; Rosenzweig 1995; Scheiner 2003). Hubbell (2001) further asserted that the relationship between diversity and area can be described using a sigmoidal curve that is asymptotic at the local scale, nearly linear at the regional scale, and exponential at the continental scale (Hubbell 2001:158).

To accommodate large numbers of observation units and to linearise relationships, natural logarithmic transformations, ln(.), are often used. We assume that the relationship between the natural logarithm of the number of tree species and the natural logarithm of area observed follows a mostly linear to asymptotic curve from the local to the global scale, which characterises species–area curves for island biogeography (MacArthur and Wilson 1967). The linearity of the relationship between the natural logarithm of the number of species and the natural logarithm of area derives from the ln–ln transformation of the equation $N = kA^z$, where N is the number of species observed for area A, and k and z are constants (Arrhenius 1921; Preston 1960); i.e., $\ln(N) = \ln(k) + z \ln(A)$.

The assumption of increasing species numbers underlying the curves is that for a generally homogenous habitat, the number of species increases with increasing plot sizes or number of plots due to small patches of heterogeneity in the spatial distribution of species. For the regional to large scale, the additional habitats are assumed to produce discrete step increases in species numbers that result in different curve shapes. This assumption is the basis for our decision to assess tree species that dominate and form forests on approximately one-third of earth's terrestrial surface and contribute almost half the terrestrial net primary biomass production (Groombridge and Jenkins 2002).

Noss (1990) includes the number of tree species among the factors that contribute to the compositional component of biodiversity. Winter et al. (2008) and Chirici et al. (2011) regard tree species diversity as a crucial component of biodiversity for reporting purposes and note that tree species is a variable that can be readily assessed by national forest inventories (NFIs). Both native and non-native tree species contribute to increasing tree species diversity, although native tree species have greater relevance for purposes of maintaining natural ecosystem dynamics and of maintaining specific diversities such as for insects, mammals and plant species, which are evolutionarily linked with the tree species (Peterken 2001).

Large-scale monitoring of forest sustainability requires ecological-based surveys that produce large databases of standardised or harmonised data. Although large-scale monitoring efforts (Fischer et al. 2009; Granke et al. 2009) have been undertaken for purposes such as assessing environmental effects on forest ecosystems, a harmonised approach to large-scale monitoring is still not available. Our approach focuses on using NFI data because NFIs are conducted in many countries and because all NFIs record tree species. The primary objective of NFIs is to produce forest planning information such as estimates of the sustainable yield and wood quality distribution of forests; thus, measures of tree diameter and tree species are common NFI variables (Winter et al. 2008; Chirici et al. 2011). Our approach uses this common NFI information to assess tree species diversity at large scales. Our intent was to demonstrate the utility of NFI data for comprehensive, large-scale monitoring to assess ecological aspects of forest sustainability (Puumalainen et al. 2003; Tomppo et al. 2010).

Our general objectives were methodological and focused on (1) determining the number of NFI plots necessary to produce acceptably precise estimates of tree species diversity and the proportion of native tree species at different geographic scales and on (2) describing the

relationships among estimates of tree species diversity for different forest categories. To accomplish both objectives, we used Monte Carlo bootstrap simulations and two distance-dependent geospatial approaches. Magnussen (2009) used Monte Carlo simulations to investigate the uncertainty of forest tree species diversity estimates using a relatively small data set; Wiegand et al. (2007) used Monte Carlo simulations to describe the spatial distribution of tree species diversity; and Miyamoto et al. (2008) used Monte Carlo simulations to estimate the sample size, in terms of number of trees, necessary to obtain accurate measures of allele richness in ash forests of *Fraxinus excelsior* L. These reports all indicate that Monte Carlo simulation techniques are appropriate for assessing the uncertainty of forest diversity estimates.

For this study, we used Monte Carlo bootstrap simulations to conduct general uncertainty assessments for tree species-based indicators of diversity using large samples of different sizes. In particular, our specific objectives focused on assessing biodiversity and were fourfold: (1) to evaluate the behaviour of species–area curves across geographic scales, (2) to determine sample sizes necessary to produce precise estimates of tree species diversity, (3) to evaluate the relationship among estimates of tree species diversity for local to large geographic scales, and (4) to determine whether the proportion of native species may be precisely estimated using sample sizes smaller than those necessary to estimate tree species diversity.

Data

Our study on the relationships among estimates of tree species diversity from local to large scales required a large data set. Such a data set was constructed by COST Action E43 *Harmonisation of National Forest Inventories in Europe: Techniques for Common Reporting*, organised under the auspices of the European program *Cooperation in Science and Technology* (COST) (COST 2009; Tomppo et al. 2010). COST Action E43 included a working group that focused on forest biodiversity (Chirici et al. 2011) and that populated a common database with NFI plot data. The database was used to test methods for harmonising estimates of forest biodiversity using NFI data from 16 regional units including three ecoprovinces of the United States of America (USA) and 13 European countries: Austria, Belgium, Czech Republic, Denmark, Finland, Germany, Ireland, Italy, Norway, Portugal, Spain, Sweden and Switzerland. The three ecoprovinces of the USA are defined by Bailey (1976): Ecoprovince 212, consisting of Lauretian mixed forest province in the north central USA in the vicinity of the Great Lakes; Ecoprovince 231, consisting of southern mixed forest province in the central region of the south-eastern USA; and Ecoprovince 242, consisting of Pacific lowland mixed forests in the western USA in parts of Oregon and Washington. Although Ecoprovinces 212 and 231 are not entirely contiguous, data for only the largest contiguous portions of these ecoprovinces were used. The database included observations of 256,513 trees representing 228 species observed on 12,536 NFI plots. Although McRoberts et al. (2009) and Gotelli and Colwell (2001) demonstrated that different plot sizes result in different diversity estimates, the different minimum NFI plot sizes (0.025–0.25 ha) are not problematic for our investigations for several reasons. First, the objective of our study was to describe the tree species diversity within our database, which covered local to large scales; the objective was not specifically to estimate tree species diversity for European countries or ecoprovinces of the USA or forest categories. Second, the large number of iterations (2,500—see below) ensures that plots of different sizes are selected in proportion to their occurrence in the database. Third, differences in plot sizes for the different regional units do not seriously affect estimates of tree species diversity. Additionally, although the minimum plot size of 670 m^2 for the American ecoprovinces is at the smaller end of the continuum of European plot sizes, estimates of tree species diversity for the American ecoprovinces are larger than for the European regions. This result can be attributed to the greater number of species in American forests than in European forests rather than to differences in plot sizes and confirms our assertion that our database is appropriate for the presented analyses. Further, we harmonised the database by using only trees with diameter at breast height (dbh) greater than 120 mm, which is the minimum dbh for NFIs whose data were included in the COST Action E43 database (Winter et al. 2008).

Methods

We used three approaches to investigate tree species diversity: a Monte Carlo simulation technique to estimate the uncertainty of estimates of biodiversity indicators and two geospatial approaches to estimate relationships among estimates of biodiversity indicators for different geographic scales. First, we constructed species–area curves using Monte Carlo simulations; second, we estimated tree species diversity for groups of regional units separated by similar distances; and third, we constructed traditional species–area curves for comparison with the curves of Hubbell (2001) and MacArthur and Wilson (1967). European plots represented in the common database were additionally assigned to classes based on European forest categories following the descriptions of the European environment agency (EEA 2006).

For the analyses, we used two tree species-related biodiversity indicators. The first indicator was *the number of tree species*, *nspe*, based on the plant list for European tree species (ICP 2010) and on the National Core Field Guide (USDA Forest Service 2007) for species in the USA. The second indicator was *the proportion of native tree species*, *pnat*. European tree species were defined as native if they were included in the EEA (2006) description of the forest categories. For European tree species included in the database but not among the EEA descriptions, we acquired additional information on their natural distributions by forest categories. Assertions that tree species were not native in a particular European forest category were carefully checked using tree species distribution maps and descriptions of Meusel et al. (1987), Meusel and Jäger (1992) and of Schütt et al. (1992). The nativeness of the species in the three American ecoprovinces was based on the expert opinion of inventory specialists who determined whether the species were indigenous to North America and, if so, whether they were found within their natural geographic ranges. A small proportion of native tree species indicates less natural forests (Winter et al. 2010) with an associated loss of biodiversity.

We used the standard deviation of the distribution of sample-based indicator estimates to describe the precision of an estimate of an indicator. The standard deviation can be estimated using bootstrap resampling implemented with Monte Carlo simulation techniques that randomly resample with replacement from the available sample data (Metropolis and Ulam 1949; Hoffman 1998; Efron and Tibshirani 1994). Efron and Tibshirani (1994) recommend at least 200 bootstrap resamples for estimating means but do not address the number of resamples necessary to obtain stable estimates of the standard deviation.

The Monte Carlo bootstrap simulations were first used to determine the number of bootstrap resamples and the number of NFI plots necessary to obtain stable estimates of the standard deviation of estimates of the two biodiversity indicators. The first set of simulations was implemented as follows. For $m = 5$, 10, 100 and 200, we combined the data for m randomly selected plots, treated the combined data as a single data set and calculated estimates of the indicators. We then randomly selected with replacement m more plots, treated the combined data for these m plots as a single second dataset and calculated estimates of the indicators. For each indicator, we now had two estimates of each indicator, which permitted calculation of the standard deviations of the estimates. We then randomly selected with replacement m more plots, treated the combined data for these m plots as a third single dataset and calculated the estimates of indicators. For each indicator, we now calculated the standard deviation of the three estimates. For each value of m, we continued this procedure in the same manner until we had randomly selected a maximum of $I = 5{,}000$ datasets where the random selection with replacement of a sample of m plots was characterised as an iteration i ($i = 1, \ldots, I$).

For each value of m, the mean value of the indicator, denoted $\hat{\mu}^{m,I}$, over I bootstrap resamples or iterations was calculated as,

$$\hat{\mu}^{m,I} = \frac{1}{I}\sum_{i=1}^{I}\hat{\mu}_i^m, \tag{1}$$

where $\hat{\mu}_i^m$ is the estimate obtained from the ith bootstrap resample. The variance of the estimate was then calculated as,

$$\mathrm{Var}(\hat{\mu})_I = \frac{1}{I-1}\sum_{i=1}^{I}\left(\hat{\mu}_i^m - \hat{\mu}^{m,I}\right)^2, \tag{2}$$

and the standard deviation was simply the square root of the variance.

For each indicator and for each value of m, we graphed the standard deviation of the indicator against the number of iterations, which ranged from 2 to 5,000, to determine the number of iterations for which the standard deviation stabilised. The graphs showed that 2,500 replications were sufficient to stabilise the estimate of the standard deviations (Fig. 1).

For the second set of Monte Carlo simulations, the number of iterations was held constant at 2,500. The entire database was divided into two subsets, one for European plots and one for American plots. This division separates the two main biogeographic regions with their different species populations (Pielou 1972). From each subset, one plot was randomly selected for each of 2,500 resampling iterations; for each iteration, the indicators were estimated;

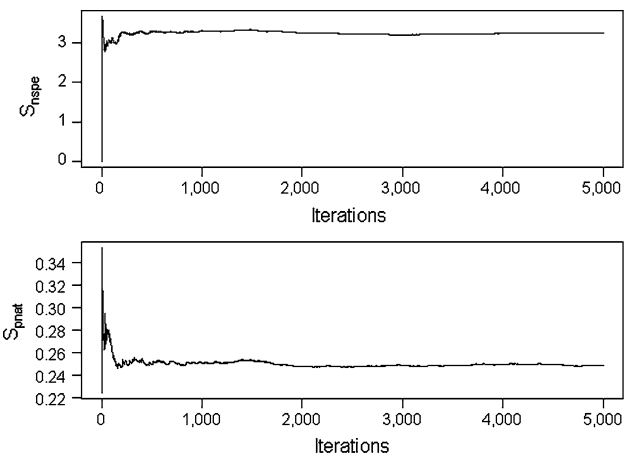

Fig. 1 Standard deviation (S) of indicator estimates (*nspe* number of tree species, *pnat* proportion of native tree species) versus number of iterations with random selection with replacement of $m = 5$ NFI plots per iteration for all plots represented in the database

and the mean and standard deviation over all iterations were calculated. Next two plots were randomly selected from each subset for each of 2,500 iterations; for each iteration, the data for the two plots were combined and the indicators were estimated for the combined dataset; and the mean and standard deviation over all iterations were calculated. This procedure was repeated for 3, 4, …, k randomly selected plots for each iteration where $k = 7{,}212$ for Europe and $k = 5{,}324$ for the USA. In addition, the European portion of the database was divided into forest category classes, and the same simulation procedure was implemented with the number of plots randomly resampled per iteration ranging from 1 to 100.

We also used two geospatial analyses that took into account distances between centroids of the regional units whose NFI plots were represented in the database. The first geospatial analyses entailed grouping the regional units with respect to the distances between them. Distances were based on the unit's centroids. Centroids for Ecoprovinces 212 and 231, which are not entirely contiguous, were calculated for only the largest portions to correspond to the areas for which data were used. The distribution of distances between the regional unit centroids showed that most regional units were within 2,000 km of the other regional units (Fig. 2). However, because of the Atlantic Ocean separating Europe and the USA, we had a data gap in the distances from approximately 3,500–6,500 km. The data for the regional units were grouped by five distance intervals with break points of 150, 1,000, 2,000 and 3,000 km. Finland and Portugal have the maximum separation distance of approximately 3,400 km among Europe countries. The maximum distance from Finland to an Ecoprovince in the USA is approximately 12,300 km to the western ecoprovince 242. The distance groups were chosen to represent a gradient from the regional to the large scale. In accordance with the general principles of spatial correlation, we assumed that species diversities for regional units separated by small distances would be more similar and smaller than the diversities for regional units that were separated by greater distances (Tobler 1970). Consequently, we expected that the average number of tree species would increase when data were grouped together for regional units separated by greater distances (Goslee 2006). This group-distance analysis produces a species–area curve that is not characterised by discrete distance steps as is the case for the second geospatial analyses.

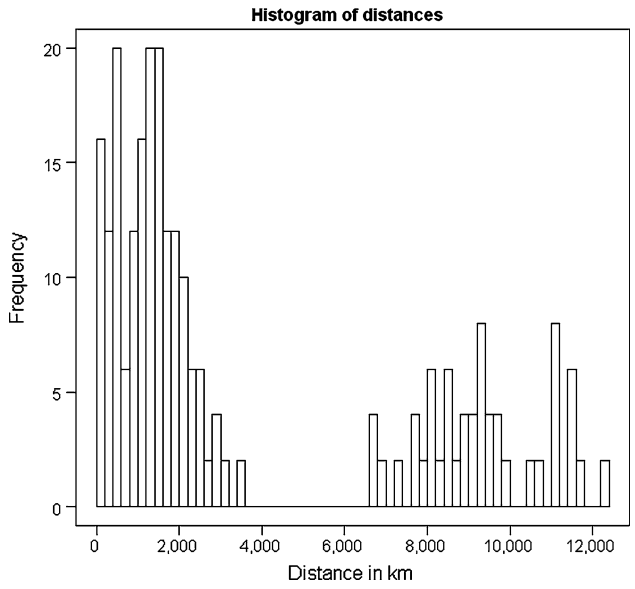

Fig. 2 Number of pairs within distance groups of 200 km between the centroids of countries and ecoprovinces that contributed NFI data to the COST Action E43 database. The total number of country and ecoprovince pairs was 256

The second geospatial analyses focused on constructing species–area curves using as area centres the centroids of four anchor regional units: Austria for central Europe, Finland for northern Europe, Portugal for southwestern Europe, and Ecoprovince 231 for the central part of the south-eastern USA. Beginning with the NFI plot data of the anchor regional unit, we constructed tree species–area curves by adding the data for regional units, depending on their increasing distances to the anchor regional unit, and using the increasing number of plots as a surrogate for area.

We used the R statistic package, version 2.10.1 (R Development Core Team 2009), for analysing the data (Bolker et al. 2009). The Monte Carlo simulations using R followed the descriptions of Robert and Casella (2010).

Results

Uncertainties in monitoring tree species diversity on large scales

For both Europe and the USA, graphs of estimates of *nspe* based on NFI data from the COST Action E43 database clearly show estimates do not approach an asymptote, meaning that the total numbers of NFI plots represented in these two subsets are insufficient to estimate tree species diversity at this large scale (Fig. 3). The two curves of mean number of species depict the decreasing uncertainty when selecting an increasing number of plots to estimate the number of tree species at large scales. However, as the number of plots increases to 2,000–4,000, estimates of mean number of tree species based on the 2,500 resampling iterations each approach asymptotes that represent environmental capacities for the number of tree species; further, these estimates are quite precise (Fig. 3, see the 2-standard deviation curves). For assessing the diversity of

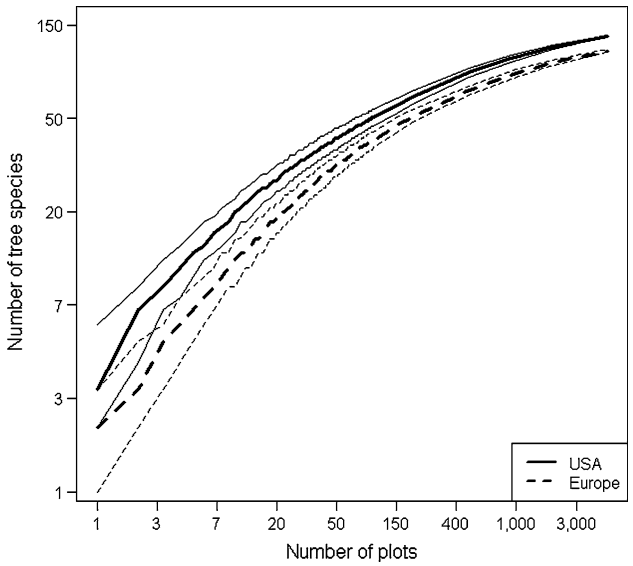

Fig. 3 Mean curves and standard deviation (s_{nspe}) intervals for the *number of tree species*, *nspe*, for NFI plot data for Europe and the USA: for Europe, *bold* and fine *continuous lines* for the mean and s_{nspe}, respectively; for the USA, *bold* and fine *dotted lines* for the mean and s_{nspe}, respectively

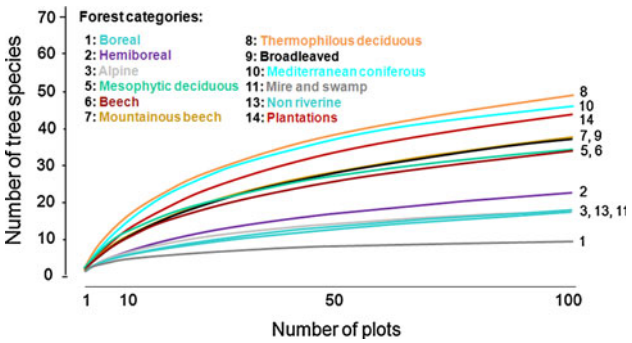

Fig. 4 Mean *numbers of tree species* versus number of NFI plots by European forest categories (EEA 2006); results for forest category 4, *Acidophilous oak forest*, and forest category 12, *Floodplain forest*, are not depicted because of fewer than 100 NFI plots in the COST Action E43 database

common and widespread species, as opposed to the tree species diversity for all tree species, which additionally includes rare species, a smaller proportion of a large database is sufficient because the probability of observing common tree species is large; approximately 10% of our database suffices.

Surprisingly, these findings also pertain to those European forest categories for which at least 100 NFI plots are represented in the database. The single exception is forest category 1, tree species-poor *boreal forests*, for which the estimated number of tree species stabilises for approximately 20 randomly selected plots, a relatively small sample size (Fig. 4). Species–area curves for forest categories 8–10 for southern Europe, particularly forest category 8, *Thermophilous deciduous forest*, and forest category 10, *Mediterranean coniferous forests*, which are rich in tree species show the least indication of approaching an asymptote. However, tree species diversity is close to stabilising using only 100 plots of species-rich forest categories. For forest category 14, *plantations*, a species–area curve would be expected to indicate continuously increasing numbers of species because of the large number of non-native species planted in the various regional units.

The proportion of native tree species, *pnat* (Fig. 5a, b), can be precisely estimated using fewer NFI plots than *the number of tree species*, *nspe* (Fig. 3). For the three American ecoprovinces, *pnat* is nearly constant at *pnat* = 1.00. For European species, as the number of plots increases, the uncertainty in estimates of *pnat* decreases. For 2,500 resampling iterations, mean *pnat* is achieved for random selections of 30 or more NFI plots. However, a single sample of 30 randomly selected plots results in a 1-standard deviation interval of [0.25, 0.55] around mean *pnat*; this interval decreases to [0.35, 0.45] for 100 randomly selected plots. For fewer than 30 plots, the slightly greater mean *pnat* and the larger standard deviation indicate that NFI plots with greater numbers of trees include more non-native trees than NFI plots with smaller numbers of trees. This result indicates that trees of non-native species are comparatively young and small and that they are increasingly planted, which is not the case for trees in older stands.

For European forest categories, random selection of 30 or more NFI plots produced precise estimates of mean *pnat* (Fig. 5b). For forest categories, $0.4 \leq pnat \leq 1.0$ shows that large-scale diversity estimates differ by forest categories. For forest category 8, *Thermophilous deciduous forest*, and forest category 10, *Mediterranean coniferous forest*, *pnat* is medium to large, which suggests that the proportion of non-native tree species does not account for the failure of the curves of *nspe* versus numbers of plots to approach the total *nspe* (Fig. 3). However, for forest category 5, *Mesophytic deciduous forest*, with *pnat* <0.5, the large number of non-native tree species partially explains why estimates of *nspe* do not approach an asymptote (Fig. 3). The occurrence of non-native tree species is constrained only by their growth potential in geographic regions corresponding to forest categories. Thus, the number of non-native tree species may be large in some forest categories, as indicated by small proportions, *pnat*, and is then difficult to estimate precisely using only 100 plots.

However, the tree species–area curve based on average *nspe* in distance groups approach asymptotes for the four smallest distance groups (Fig. 6). The largest distance group combines data for the two most distant European countries of Finland and Portugal and the ecoprovinces of

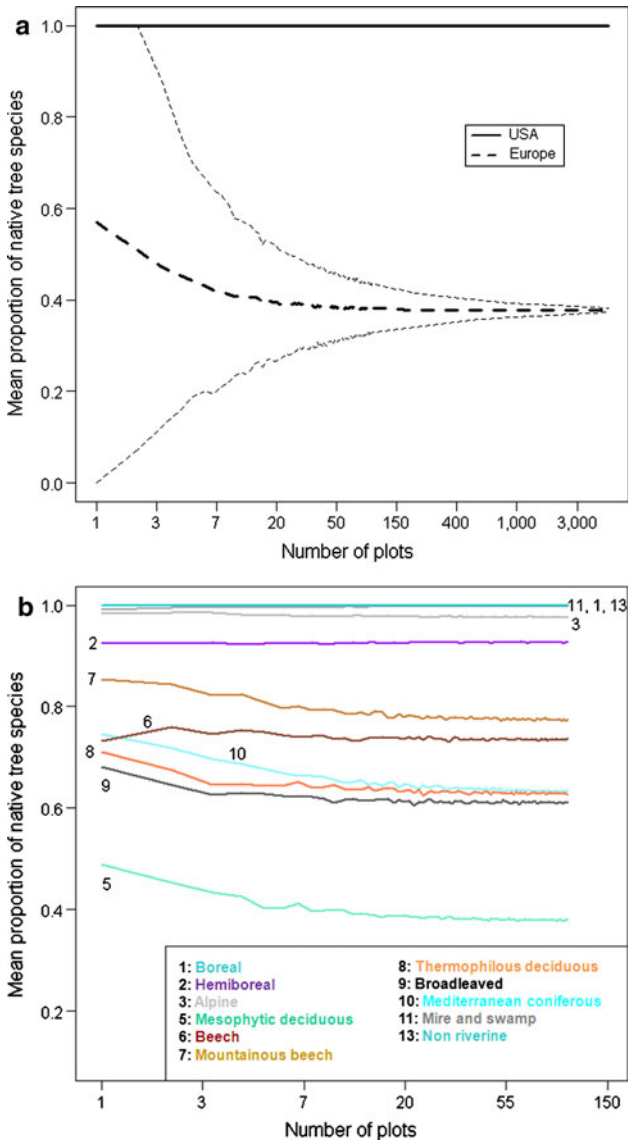

Fig. 5 *Proportions of native tree species*: **a** means and standard deviations, S_{pnat}, for Europe and the USA according to an increasing number of plots (1 to the total number of plots for Europe (7,212) and the USA (5,324)). The mean and standard deviation is estimated from the results of 2,500 resampling iterations for each of the plot numbers (1-total); **b** means for European forest categories where forest category *4*, *Acidophilous oak forest*, and forest category *12*, *Floodplain forest*, are not depicted because of fewer than 100 NFI plots in the COST Action E43 database. Forest category *14*, *plantations*, is not depicted because without coordinates of the plots, the nativeness of a tree species cannot be assessed

the USA; thus, this group is the most heterogeneous. Nevertheless, the general shape of the curve is similar to the linear species–area curves described for island biogeography (MacArthur and Wilson 1967) with a slight sigmoidal shape as per Hubbell's curve. The tree species–area curve is asymptotic for the European data and is nearly linear for the entire data set.

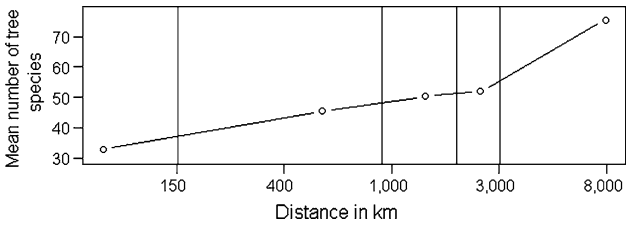

Fig. 6 Species–distance curve of the mean *number of tree species* for five distance groups with the breaks of 150, 1,000, 2,000 and 3,000 km

Species area–curve for a discrete distance gradient

The combination of data for scales from local to large scale (Fig. 7) reveals three aspects of the relationship between *nspe* and the associated area represented by the number of plots. First, as the number of plots increases within the anchor regional unit (e.g., 658 Austrian plots, Fig. 7a), the curves clearly approach asymptotes. At the large scale, the curves remain asymptotic for Portugal (Fig. 7c) as an anchor or become nearly linear for Austria, Finland (Fig. 7a, b) and Ecoprovince 231 (Fig. 7d) as anchors. Second, *nspe* is quite linear when data for regional units are aggregated to a continental scale. Third, it is obvious that the 5,000 NFI plots from the three American ecoprovinces represented in the COST Action E43 database are insufficient to estimate *nspe* for these three ecoprovinces. Thus, the curves starting with Austria, Finland and Portugal as anchor countries show large increases in *nspe* without approaching an asymptote.

The tree species–area curves for different anchor countries are of different shapes (Fig. 8) using ln–ln transformations. For Austria and Portugal as anchor countries, the increase in the numbers of species results in similar curve shapes that are within the envelope of the curves obtained when using Finland and American Ecoprovince 231 as anchors. The four curves cover the range of possibilities of species–area curve shapes described in the introduction.

Discussion

Uncertainties of estimates of biodiversity indicators at large geographic scales

Estimates obtained from efforts to monitor biodiversity and from scenarios for assessing the loss of biodiversity are subject to uncertainty (for vascular plants, see van Vuuren et al. 2006). In the regions of our study, we have a quite complete overview on the diversity of plant species,

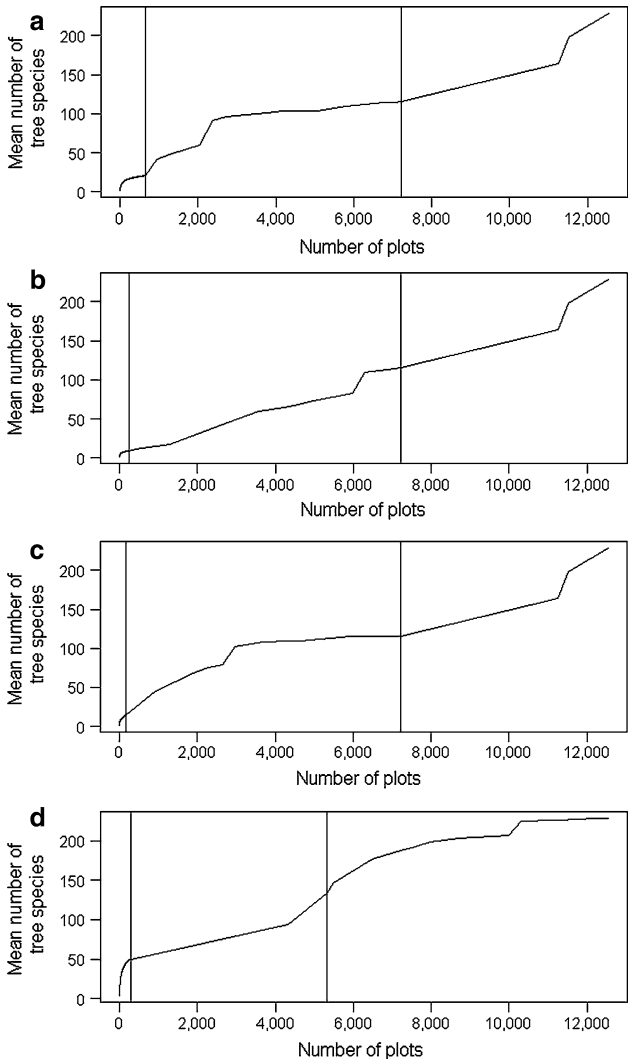

Fig. 7 Tree species–area curves for different anchor countries (**a** Austria, **b** Finland and **c** Portugal) and the Ecoprovince 231 of USA (**d**); *solid vertical lines* indicate transitions from (1) the local to regional (first neighbouring countries or ecoprovinces) level and (2) continent to continent level (*second vertical line*)

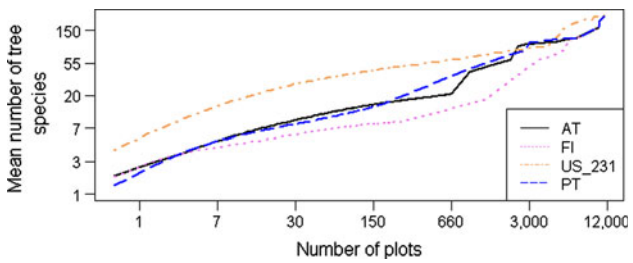

Fig. 8 Tree species–area curves (ln–ln data transformation) constructed for different anchor countries: Austria (*AT*), Finland (*FI*), Ecoprovince 231 of the USA (US_321) and Portugal (*PT*); the order of the increasing plot numbers is, first, all plots in anchor countries or ecoprovince and, second, averaged numbers of tree species in countries with increasing distances from the anchor country

especially of tree species. However, comparisons of the number of tree species in our database to the total number of tree species in the regional units are not possible. The number of tree species is largely dependent on factors such as species definitions, definitional distinctions between tree and shrub species, and decisions as to whether tree species in botanical gardens, experimental forests and rural areas are included. Consequently, reliable estimates of total numbers of tree species in our regional units or within the forest categories are not available.

Uncertainties associated with monitoring at large scales result from a combination of lack of knowledge concerning biodiversity stressors influencing native tree species distributions, inaccuracy in field work and data management, and most notably inaccuracy and differences in methods used to estimate biodiversity indicators at international scales (Hoffmann-Kroll et al. 2003; Chirici et al. 2011). Our study focused on the uncertainty of large-scale estimates of tree species-related indicators using NFI data. The results of our Monte Carlo bootstrap simulations clearly indicate that comprehensive monitoring of tree species diversity at large scales requires an extremely large database. We found that combining the NFI data for a large number of European countries and/or some American ecoprovinces at least produces precise estimates of large-scale tree species diversity indicators. However, the nature of the tree species abundance distributions, with small numbers of common and widespread species and large numbers of specialists in small niches, which is typical for most species groups, makes samples sizes to observe all species prohibitive in many cases (Chao et al. 2009). Thus, efforts have focused on sample-based assessments rather than complete species enumeration, which has already been suggested for small-scale biodiversity (Longino et al. 2002).

Several limitations associated with the analyses are clear. First, the different NFI designs with different plot sizes may result in different accuracies of tree species composition based on data for single plots. However, the large number of plots considered is assumed to overcome this limitation. Second, the minimum dbh of 120 mm used for harmonising the tree data may limit some results. For example, some tree species typically remain in the understory where they grow only slowly and seldom greater than 120 mm dbh; thus, a smaller minimum dbh for NFI tree records is highly recommended. Finally, tree species, especially for rare and introduced tree species that cannot be easily identified, should not be recorded using broad categories such as "other oaks" or "other coniferous tree species". NFIs that use this practice will tend to underestimate tree species diversity. Further, tree species diversity analyses may better be restricted to contiguous mainlands rather than to continents. Analyses that span different

continents with distinct evolutionary species histories produce different tree species compositions which, when combined, produce bimodal species distributions.

Relationships between number of species and area

The Monte Carlo simulations and both geospatial methods based on distances show that the relationship between *nspe* and area are not fully or adequately described by any particular existing model for species–area curves. Although the curves are clearly different in their forms, they tend to be linear when using ln–ln transformations. The shapes of curves depicting the relationship between the *nspe* and corresponding area depend more on the geographic scale of the data and on the choice of the anchor regional unit than on factors such as species group or features of the island biogeography expressed by island size and distance to the next continent (Kalmar and Currie 2006). In this context, the order of the regional units with respect to their distances to the anchor regional unit is determined by the selected anchor regional units in different geographic regions. Some features of the three models for species–area curves, Arrhenius' (1921) power model, Hubbell's (2001) sigmoidal model and the asymptotic model are reflected in the species–area curves obtained using the COST Action E43 tree species data. The exponential model proposed by Gleason (1922) does not adequately describe the species–area curves because none of our curves is exponential after ln–ln transformation. In our study, we modified and combined the species–area curve types described by Scheiner (2003). We analysed non-continuous data using a combination of partially smooth curves and curves with discrete, discontinuous steps. Scheiner (2003) separates the curves at the locations of these steps. The combination of different approaches is admissible because the smooth curve shapes for the anchor regional unit represent averaged subdivisions of individual regional units at the larger scales. Local diversity in our study generally conforms to an asymptotic curve. The asymptotic shape of the local to regional species–area curve is confirmed theoretically (Harte et al. 2009) and by several studies on different species groups (Rosenzweig 1995).

Our results show that it is generally not possible to define smooth, continuous transitions from local to large scales. And our results confirm the statement of Whittaker et al. (2001) regarding the impossibility of constructing a single, simple, universally applicable model that describes relationships between number of species and area. This result is in contrast to the work of Harte et al. (2009) who attempted to develop such a universal model for species–area curves.

Conclusions

Three conclusions may be drawn from our study. First, the Monte Carlo resampling simulations and the two geospatial analyses produced valid species–area curves and estimates of uncertainty and are appropriate for use with other species groups. Second, precise estimates of European tree species diversity based on NFI data will not be possible with any monitoring approach that does not include data representing nearly all of Europe obtained using large sampling intensities. Third, for the total number of tree species, estimates of the uncertainties of estimates reflect the unfavourable state of global biodiversity monitoring of species groups. The numbers of species observed for the European forest categories and the American ecoprovinces are small compared with tree species-rich forest regions; nevertheless, even with a database as large as that used for this study, the total number of tree species cannot be adequately estimated. Consequently, biodiversity monitoring of global forests is currently a highly uncertain enterprise.

Acknowledgments Special thanks to Erkki Tomppo, the chairperson of the COST Action E43 Management Committee. Without his excellent leadership, this paper would not have been possible. We also thank Annemarie Bastrup-Birk, Marco Marchetti and especially Gherardo Chirici for their excellent coordination and leadership of Working Group 3 of COST Action E43. We thank all the countries that authorised provision of NFI data for the common database: Anna-Lena Axelsson for Sweden, Annemarie Bastrup-Birk and Vivian Kvist Johannssen for Denmark, Christine Sanchez for Belgium, Elmar Hauk for Austria, Iciar Alberdi for Spain, Jana Beranova for Czech Republic, Jan-Erik Nilsen for Norway, John Redmond for Ireland, Paulo Godinho-Ferreira for Portugal, Tiina Tonteri for Finland, Urs-Beat Brändli for Switzerland and Vittorio Tosi for Italy. Thanks to Hans-Joachim Klemmt for the first version of the R script on the Monte Carlo simulation. Thanks to Lisa Mahal, University of Nevada, Las Vegas, for GIS assistance.

References

Arrhenius O (1921) Species and Area. J Ecol 9:95–99

Bailey RG (1976) Ecoregions of the United States (map). USDA Forest Service, Intermountain Region, Ogden

Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B (2009) gplots: various R programming tools for plotting data, R package version 2.7.4. http://CRAN.R-project.org/package=gplots

Boutin S, Haughland DL, Schieck J, Herbers J, Bayne E (2009) A new approach to forest biodiversity monitoring in Canada. Forest Ecol Manag 258(S1):S168–S175

Brockway DG (1998) Forest plant diversity at local and landscape scales in the Cascade Mountains of southwestern Washington. Forest Ecol Manag 109:323–341

Convention on Biological Diversity (CBD) (1992) Convention on biological diversity. Secretariat of the convention on biological diversity, United Nations environment programme, p 83

Chao A, Colwell RK, Lin CW, Gotelli NJ (2009) Sufficient sampling for asymptotic minimum species richness estimators. Ecology 90:1125–1133

Chirici G, Winter S, McRoberts RE (2011) Contribution of national forest inventories for forest biodiversity assessment. Springer, Dordrecht, Netherlands, p 622

COST (2009) European Cooperation in the field of Scientific and Technical Research. Available at: http://www.cost.esf.org/. Last accessed Jan 2010

Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. Chapmann/Hall, New York, p 436

European Environment Agency (EEA) (2006) European forest types. Categories and types for sustainable forest management and reporting, p 111. EEA technical report no. 9/2006

Fischer R, Granke O, Chirici G, Meyer P, Seidling W, Stofer S, Corona P, Marchetti M, Travaglini D (2009) Background, main results and conclusions form a test phase for biodiversity assessments on intensive forest monitoring plots in Europe. iForest-Biogeosci 2:58–65. doi:10.3832/ifor0493-002

Gleason HA (1922) On the relation between species and area. Ecology 3:158–162

Goslee SC (2006) Behavior of vegetation sampling methods in the presence of spatial autocorrelation. Plant Ecol 187 (2):203–212

Gotelli NJ, Colwell RK (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. Ecol Lett 4:379–391

Granke O, Kenter B, Kriebitzsch WU, Köhl M, Köhler R, Olschofsky K (2009) Biodiversity assessment in forests—from genetic diversity to landscape diversity. iForest-Biogeosci 2:1–3. doi: 10.3832/ifor0493-002

Groombridge B, Jenkins MD (2002) World Atlas of biodiversity. University of California Press, Berkeley and Los Angeles, California, p 340

Harte J, Smith AB, Storch D (2009) Biodiversity scales from plots to biomes with a universal species-area curve. Ecol Lett 12:789–797

Hoffman P (1998) The man who loved only numbers. The story of Paul Erdos and the search for mathematical truth. Hyperion, New York, pp 238–239

Hoffmann-Kroll R, Schäfer D, Seibel S (2003) Landscape indicators from ecological area sampling in Germany. Agric Ecosyst Environ 98:363–370

Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, p 373

ICP (2010) International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests in co-operation with the European Commission: http://www.icp-forests.org/. Last accessed Jan 2010

Kalmar A, Currie DJ (2006) A global model of island biogeography. Global Ecol Biogeogr 15(1):72–81

Longino JT, Coddington J, Colwell RK (2002) The ant fauna of a tropical rain forest: estimating species richness three different ways. Ecology 86:1143–1153

MacArthur RH, Wilson EO (1967) The theory of island biogeography. Princeton Landmarks in biology. Princeton University Press, Princeton, p 224

Magnussen S (2009) A balanced repeated replication estimator of sampling variance for apparent and predicted species richness. Forest Sci 55(3):189–200

MCPFE (1997) Work-programme on the conservation and enhancement of biological and landscape biodiversity in forest ecosystems. Adopted at expert-level by the third meeting of the executive Bureau of the Pan-European biological and landscape diversity strategy 20–21 Nov 1997, Geneva, Switzerland

MCPFE (2003) Vienna declaration and Vienna resolutions adopted at the fourth ministerial conference on the protection of forests in Europe, 28–30 Apr 2003, Vienna Austria

McRoberts RE, Tomppo E, Schadauer K, Vidal C, Ståhl G, Chirici G, Lanz A, Cienciala E, Winter S, Smith WB (2009) Harmonizing National Forest Inventories. J For 107:179–187

Metropolis N, Ulam S (1949) The Monte Carlo method. J Am Stat Assoc 44:335–341

Meusel H, Jäger E (1992) Vergleichende Chorologie der zentraleuropäischen Flora. Maps, Fischer Verlag, Jena, p 421

Meusel H, Jäger E, Rauschert S, Weinert E (1987) Vergleichende Chorologie der zentraleuropäischen Flora. Fischer Verlag, Jena, p 688

Miyamoto N, Fernandes-Manjarres JF, Morand-Prieur ME, Bertolino P, Frascaria-Lacoste N (2008) What sampling is needed for reliable estimations of genetic diversity in *Fraxinus excelsior* L. (Oleaceae)? Ann Forest Sci 65(4):8, Article 403

Montréal Process (2009) Criteria and indicators for the conservation and sustainable management of temperate and boreal forests. www.rinya.maff.go.jp. Last accessed 8 Feb 2010

Noss R (1990) Indicators for monitoring biodiversity: a hierarchical approach. Conserv Biol 4:355–364

Peterken GF (2001) Ecological effects of introduced tree species in Britain. Forest Ecol Manag 141(1–2):31–42

Pielou EC (1972) Niche width and niche overlap: a method for measuring them. Ecology 53(4):687–692

Preston FW (1960) Time and space and the variation of species. Ecology 41(4):611–627

Puumalainen J, Kennedy P, Foving S (2003) Monitoring forest biodiversity: a European perspective with reference to temperate and boreal forest zone. J Environ Manag 67:5–14

R Development Core Team (2009) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Robert C, Casella G (2010) Introducing Monte Carlo methods with R. Springer, Heidelberg

Rosenzweig ML (1995) Species diversity in space and time. Cambridge University Press, Cambridge, p 436

Scheiner SM (2003) Six types of species-area curves. Global Ecol Biogeogr 12:441–447

Schoener TW (1976). The species-area relation within archipelagos: models and evidence from island land birds. In: Proceedings of the 16th international ornithological conference, Australian Academy of Science, Canberra, pp 629–642

Schütt P, Schuck HJ, Stimm B (1992). Lexikon der Forstbotanik. Ecomed, Landsberg am Lech, p 581

Tobler W (1970) A computer movie simulation urban growth in the Detroit regions. Econ Geogr 46(2):234–240

Tomppo E, Gschwantner T, Lawrence M, McRoberts RE (2010) National forest inventories: pathways for common reporting (eds). Springer, Dordrecht, Netherlands, p 670

USDA Forest Service (2007) US forest inventory and analysis, National core field guide. Volume 1: field data collection procedures for Phase 2 plot, Version 4.0, p 224. US Department of Agriculture. Available at: http://fia.fs.fed.us/library/field-guides-methods-proc/. Last accessed 22 Mar 2010

Van Vuuren DP, Sala OE, Pereira HM (2006) The future of vascular plant diversity under for global scenarios. Ecol Soc 11(2):19, Article 25

Watson HC (1859) Cybele Britannica, or British plants and their geographical relations. Longman & Company, London

Whittaker RH (1972) Evolution and measurement of species diversity. Taxon 21:213–251

Whittaker RH (1975) Communities and ecosystems, 2nd edn. Macmillan, New York, p 385

Whittaker RH (1977) Evolution of species diversity in land communities. In: Hecht MK, Steere WC, Wallace B (eds) Evolutionary biology 10. Plenum Press, London, pp 325–335

Whittaker RJ, Willis KJ, Field R (2001) Scale and species richness: towards a general, hierarchical theory of species diversity. J Biogeogr 28:453–470

Wiegand T, Gunatilleke S, Gunatilleke N (2007) Species associations in a heterogeneous Sri Lankan dipterocarp forest. Am Nat 170(4):E77–E95

Winter S, Chirici G, McRoberts RE, Hauk E, Tomppo E (2008) Possibilities for harmonising national forest inventory data for use in forest biodiversity assessments. Forestry 81:33–44

Winter S, Fischer HS, Fischer A (2010) Relative quantitative reference approach on naturalness assessments. Forest Ecol Manag 259:1624–1632