Biodiversity and Conservation (2023) 32:1019–1046
https://doi.org/10.1007/s10531-022-02536-0

ORIGINAL RESEARCH

Insights in forest structural diversity indicators with machine learning: what is indicated?

J. S. Ette[1] · T. Ritter[2] · S. Vospernik[2]

Received: 10 May 2022 / Accepted: 31 December 2022 / Published online: 6 January 2023

Abstract
Indicator choice is a crucial step in biodiversity assessments. Forest inventories have the potential to overcome data deficits for biodiversity monitoring on large spatial scales which is fundamental to reach biodiversity policy targets. Structural diversity indicators were taken from information theory to describe forest spatial heterogeneity. Their indicative value for forest stand variables is largely unknown. This case study explores these indicator–indicandum relationships in a lowland, European beech (*Fagus sylvatica*) dominated forest in Austria, Central Europe. We employed five indicators as surrogates for structural diversity which is an important part of forest biodiversity i.e., Clark & Evans-, Shannon, Stand Density, Diameter Differentiation Index, and Crown Competition factor. The indicators are evaluated by machine learning, to detect statistic inter-correlation in an indicator set and the relationship to twenty explanatory stand variables and five variable groups on a landscape scale. Using the R packages *randomForest*, *VSURF*, and *randomForest Explainer*, 1555 sample plots are considered in fifteen models. The model outcome is decisively impacted by the type and number of explanatory variables tested. Relationships to interval-scaled, common stand characteristics can be assessed most effectively. Variables of 'stand age & density' are disproportionally indicated by our indicator set while other forest stand characteristics relevant to biodiversity are neglected. Within the indicator set, pronounced inter-correlation is detected. The Shannon Index indicates the overall highest, the Stand Density Index the lowest number of stand characteristics. Machine learning proves to be a useful tool to overcome knowledge gaps and provides additional insights in indicator–indicandum relationships of structural diversity indicators.

Keywords Biodiversity assessments · Biodiversity indicator choice · European beech forests · Forest inventories · Indicator–indicandum relationships · R randomForest

Communicated by Neil Brummitt.

✉ S. Vospernik
Sonja.Vospernik@boku.ac.at

1 Department of Forest Growth, Silviculture & Genetics, Austrian Federal Research Centre for Forests, Vienna, Austria

2 Department of Forest and Soil Sciences, Institute of Forest Growth, University of Natural Resources and Life Sciences, Peter-Jordan-Straße 82, 1190 Vienna, Austria

Introduction

The rapid rate of biodiversity loss is an emerging public concern. There is high scientific evidence for a positive relationship between the loss of biodiversity and the decline of forest ecosystem services (Hooper et al. 2005; Balvanera et al. 2006; Isbell et al. 2011; Mace et al. 2012; Gamfeldt et al. 2013). Biodiversity loss threatens the provision of ecosystem services at an accelerating rate and erodes the foundation of humanity (IPBES 2019).

The main drivers of extinction and decline are of anthropogenic origin (Sala et al. 2000; Newbold et al. 2015). Forest degradation, fragmentation, and loss as side effects of human economic activities already caused severe biodiversity losses (Newbold et al. 2015; FAO 2020). Globally, extinction rates are being one hundred to one thousand times greater than the natural baselines (Ceballos et al. 2010, 2015). This trend is expected to continue globally (Keenan et al. 2015; Newbold et al. 2015).

Acknowledging the importance of biodiversity, numerous measures in policy, public, and sciences have been taken to halt biodiversity loss. Major global initiatives are the Convention on Biological Diversity (est. 1992), the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (est. 2012), and the Sustainable Development Goals (est. 2016). At the European level, the Ministerial Conference on the Protection of Forests in Europe (est. 1990), the Streamlining European Biodiversity Indicators Initiative (est. 2005), the EU Biodiversity Strategy (est. 2011), and the European Green Deal (est. 2019) were initiated. About 14.4 billion USD was spent globally from 1992 to 2003 to halt biodiversity loss (Waldron et al. 2017). Although, the rate of biodiversity decline was below the expected decline, strategic aims to control biodiversity loss are never met (CBD 2014; Tittensor et al. 2014).

One of the reasons for environmental policy implementation gaps may be the lack of effective biodiversity monitoring systems (Pereira et al. 2012; CBD 2018; Ette and Geburek 2021). Biodiversity indicators play a crucial role in assessing biodiversity and were established in large numbers (Lindenmayer et al. 2000; Larsson et al. 2001; Chirici et al. 2011). Nonetheless, biodiversity indicators are still criticized for poor indicator–indicandum relationships (Ferris and Humphrey 1999; Margules et al. 2002; Duelli and Obrist 2003; Gao et al. 2015). Following the definition of Heink and Kowarik (2010) an indicator is of major relevance for a given issue, e.g., assessment of a certain impact on conservation policy, while an indicandum is the phenomenon indicated.

Indicators for biodiversity are considered to be more useful the more precise the correlation between indicator and indicandum is known (Heink and Kowarik 2010). Scientists, policymakers, and forest managers are facing severe knowledge gaps while having to decide which and how to choose and aggregate biodiversity indicators (Yoccoz et al. 2001; McElhinny et al. 2005; Katzner et al. 2007; Jones et al. 2011). On large spatial and temporal scales, the availability of reliable data sets is another limiting factor for biodiversity monitoring (Purvis and Hector 2000; Heym et al. 2021). Therefore, there is no forest biodiversity monitoring approach internationally established or accepted yet (CBD 2018).

Due to a lack of consistent correlations, indicator species concepts have not been successful (Margules et al. 2002; Duelli and Obrist 2003). Structural diversity indicators reflect potential habitat variability, niche differentiation, structural complexity (Heym et al. 2021), and sources of forest biodiversity (McElhinny et al. 2005) e.g., for umbrella species (Müller et al. 2009) and bird species (MacArthur and MacArthur 1961). There is broad scientific evidence for positive relationships between measures of forest structural variety and elements of biodiversity (Begon et al. 1996; McNally et al. 2001; Winter et al. 2008; Motz et al. 2010).

Forest inventories have a potential to overcome data deficits on large scales (Chirici et al. 2011; Corona et al. 2011; Storch et al. 2018). Major advantages of inventory-based biodiversity assessments are the repeated measurements which reflect temporal changes (Heym et al. 2021) with low additional costs (Corona et al. 2003, 2011) for a high number of attributes, forest types, sample sizes, and scales (Storch et al. 2018; Heym et al. 2021). In the long term, changes in biodiversity can even be related to forest management practices (Storch et al. 2018) which makes it highly reasonable to choose indicators based on forest inventory data. Handling knowledge gaps in choice and aggregation of biodiversity indicators by machine learning approaches has already been explored in permanent grassland and freshwater ecosystems (Gallardo et al. 2011; Plantureux et al. 2011).

Our case study examines the potential of this approach for forest ecosystems. In line with Noss (1990), and McElhinny et al. (2005), it focusses on tree species composition and forest structure in a surrogate approach (Olsgard et al. 2003). Scientifically well-established metrics of structural diversity relevant to forest biodiversity are applied. Although the relationship to the indicandum may not be fully understood yet, we will refer to these metrics as 'indicators' in the following.

Our goal is to promote the applicability of forest inventory-based diversity indicators by precising indicator–indicandum relationships through machine learning. Following Pretzsch (2002), the comprehensive indicator set considers horizontal distribution, tree species diversity, Stand Density and stand differentiation. Machine learning is applied to forest inventory data on a landscape scale in an unmanaged, lowland, European beech (*Fagus sylvatica* L.) dominated forest in Austria to answer the following research questions: (1) Which levels of structural diversity can be found in the unmanaged core areas of the Biosphere Reserve Vienna Woods (BR)? (2) Which stand characteristics are indicated by single structural diversity indicators? (3) Which stand characteristics are indicated or neglected by a comprehensive indicator set? (4) How strong is the intercorrelation in an indicator set?

The hypotheses of this study are that (1) machine learning as an integral part of artificial intelligence is an effective way to gain new insights in indicator–indicandum relationships in forests and (2) some stand characteristics relevant to forest biodiversity are indicated disproportionally in comprehensive indicator sets (in sense of Pretzsch 2002), while others are neglected.

Material

Study area

The case study focuses on the core areas of the Biosphere Reserve Vienna Woods in East Austria, Central Europe (48° 5′ N, 15° 54′ E). The BR Vienna Woods has an area size of 105.000 ha and was established in 2005. The study area is located in the transition zone between the Vienna Basin and the Northern Limestone Alps. The 37 core areas (5.400 ha) under strict nature protection and without forest management are scattered across the Biosphere Reserve (Fig. 1, ESM1). The dominant tree species are European beech (*Fagus sylvatica)* 57%, oak (*Quercus* spp.; *Q. robur*, *Q. petrea*, *Q. cerris*) 22%,

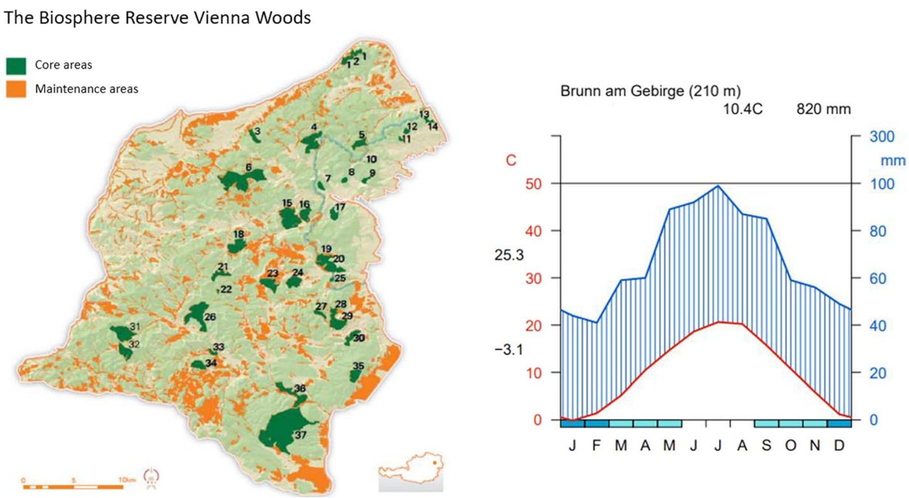

Fig. 1 Map of study area (BFW 2011) & study climate. The study is conducted in the scattered core areas of the Biosphere Reserve Vienna Woods, located in East Austria, Central Europe. Mean monthly precipitation of the climate station "Brunn im Gebirge" ranges between 41 and 99 mm. Mean monthly temperatures are between − 0.1 °C and + 20.8 °C (EHYD 2021)

hornbeam (*Carpinus betulus*) 11%, ash (*Fraxinus excelsior*) 2%, birch (*Betula pendula*) 2%, larch (*Larix decidua*) 2%, and pine (*Pinus sylvestris*) 2% (BR Vienna Woods Management 2011).

Due to beneficial climatic conditions along the Vienna Thermal Line, the landscape was intensely used for centuries for transportation, settlement, agriculture, and forest management (Schachinger 1934). Historical forest management was favoring oak, black pine, and wild fruit tree species targeting firewood, game, resin, wild fruits, and acorns (Schachinger 1934). The centrally located climate station in "Brunn im Gebirge" shows the highest average monthly precipitation in July (99 mm) and the lowest in February (41 mm). Mean monthly temperatures range between − 0.1 °C in January and + 20.8 °C in July (Fig. 1). Hydrographic examinations in the Biosphere Reserve show, that annual precipitation amount can diverge up to three times on small spatial scales (EHYD 2021). The Eastern parts of the BR are under Pannonian climate, while the North-Western parts are dominated by Atlantic climate. From a geological point of view, the area under survey can decisively be distinguished in flysch and limestone bedrock. Due to heterogeneity in terms of soil, bedrock, precipitation, and topography, the BR Vienna Woods is ecologically highly diverse. About a quarter of the 125 forest types of Austria (Mucina et al. 1993) occur in the BR.

Core area monitoring

The core area monitoring of the BR Vienna Woods consists of 1555 permanent sample plots in the 37 unmanaged core areas. Since 2007, updated field data is available in a 10 year interval. Depending on the core area size, variable grid spacing guarantees a recording accuracy of ± 10% of the living standing volume. For more details, please see the field work manual (Posch et al. 2008), the monitoring results published (BR Vienna Woods

Management 2011), and the core area overview (ESM1). Our study considers data from the first inventory period (2008–2010).

Tree species and growing stock volume

Sample trees were collected using angle count sampling (synonym: relascope sampling, Bitterlich sampling) with basal area factor $z = 4m^2\ ha^{-1}$ (Bitterlich 1984). Angle count sampling, which is commonly used in large scale forest inventories (e.g., Gabler and Schadauer 2007), is a variable radius sampling technique, with inclusion probabilities proportional to the trees' basal area. Trees are recorded according to the relation of stem diameter and distance to a central inventory point (Heym et al. 2021). Tree diameter at breast height (dbh) at 1.3 m above ground was measured for all trees in the angle count sample using a caliper. Additionally, tree height of every basal area median tree was measured per tree species and sample plot. In any case of tree top break, tree heights were additionally measured. Heights of all other trees in the sample were estimated using the basal area median tree heights and unified height curves of the Austrian National Forest Inventory (Gabler and Schadauer 2007).

Nearest neighboring tree and forest spatial structure

For each tree in the angle count sample, horizontal distance to the nearest neighboring tree was measured and recorded together with tree species and dbh of the nearest neighbor. A diameter threshold of ≥ 10 cm was applied.

Standing and lying dead wood

To estimate standing dead wood volumes, tree height and dbh of all standing dead wood within the angle count sampling (Bitterlich 1984) was measured. In addition, lying deadwood was recorded using fixed radius circular sample plots (horizontal radius $r = 8$ m) with 20 cm diameter threshold. Depending on diameter at the midpoint (dm), two different cubing tables were used to calculate the individual wood volume for objects of (i) 20 cm $<$ dm $\leq$ 50 cm ($vol_{20-50\,\mathrm{cm\,dm}}$) and (ii) dm $>$ 50 cm ($vol_{>50\,\mathrm{cm\,dm}}$). These single cubations were added up per sample point in both categories, yielding the total volume of lying deadwood with dm $>$ 20 cm ($vol_{>20\,\mathrm{cm\,dm}}$). The total volume of lying dead wood with dm $>$ 5 cm ($vol_{>5\,\mathrm{cm\,dm}}$) was deviated from this value by applying a bridging function (Eq. 1) for natural, beech dominated forests following Christensen et al. (2005):

$$vol_{>5\,\mathrm{cm\,dm}} = vol_{>20\,\mathrm{cm\,dm}} \times (0.0279 \times threshold\ dm\,[20\,\mathrm{cm}] + 0.8301) = 1.3881 \times vol_{>20\,\mathrm{cm\,dm}} \tag{1}$$

Natural regeneration

At each sample point, young trees between 10 and 130 cm height were recorded on an area of 12.5 m^2. The last year's browsing damage on leading shoots by ungulates was documented binary (browsed/not browsed).

Soil monitoring

Information about soils in the core areas is available from the BR Vienna Woods soil monitoring which was completed in 2012. Soil samples were analyzed in the laboratory by the Austrian Federal Research Centre for Forests. Every fourth sample plot of the core area monitoring was inventoried. At those 422 sample plots, bedrock, geological unit (flysch and limestone forest), soil type, humus type, and soil water balance were surveyed.

Structural diversity indicators

For a reliable assessment of structural diversity and biodiversity, it is necessary to consider comprehensive indicator sets (Pretzsch 2002; LaRue et al. 2019). This case study uses a surrogate approach (Olsgard et al. 2003). In order to assess structural forest diversity, five structural diversity indicators (Table 1) are evaluated in a comprehensive indicator set following Pretzsch (2002). Two indicators of Stand Density are chosen with the purpose to study the effect of indicator choice on indicator correlation and indicative values of comprehensive indicator sets.

The *Clark & Evans-Index (C & E)* describes the aggregation of horizontal tree distribution which is calculated by the quotient of the observed to the expected distance between neighboring trees assuming Poisson distribution (Clark and Evans 1954). The *Shannon Index (H´)* indicates the diversity of tree species and their relative abundances in a species mixture (Shannon and Weaver 1949). The *Stand Density Index (SDI)* displays the allometric relationship between quadratic mean diameter and stem density (Reineke 1933; Pretzsch 2002). The *Crown Competition factor (CCF)* as a second Stand Density indication is a relative measure of competitive pressure in crown space describing the ratio of area size and crown canopy area (Krajicek et al. 1961). The *Diameter Differentiation Index (Diff)* reveals distance-dependent structural diversity and quantifies the heterogeneity of plant stands (Füldner 1995). The choice of indicators relevant to biodiversity needs to be legitimated (Heink and Kowarik 2010). Scientific evidence for the expected relation between the structural diversity metric (indicator) and certain aspects of forest biodiversity (indicandum) in order to establish a comprehensive biodiversity indicator set is provided in Table 2.

Methods

Explanatory variables

We apply machine learning on forest inventory data to gain new insights in indicator–indicandum relationships. Twenty stand characteristics are reviewed as potential explanatory variables in ten random forests models. These variables can be grouped into five categories: (i) 'age & density', (ii) 'vertical structure', (iii) 'forest site', (iv) 'game impact', and (v) 'soil & bedrock' (Table 3). The explanatory variables tested were chosen from monitoring data available and based on literature reviews (e.g., McElhinny et al. 2005; Gao et al. 2015; Storch et al. 2018). In this case study, species distribution maps (bats, birds, amphibians, snails, insects, higher plants, mosses, lichens, and fungi)

Table 1 Structural diversity indicator set

Structural aspect	Structural diversity indicator	Formula	References
Horizontal distribution	Clark & Evans-Index	$\mathrm{C\&E} = \frac{\left(\frac{\sum_{i=1}^{z} n_{rep*r_a}}{\sum_{i=1}^{z} n_{rep}}\right)}{0.5\cdot\sqrt{\frac{10{,}000m^2}{N_{obs}}}}$	Clark and Evans (1954)
Species diversity	Shannon Index	$\mathrm{H}' = -\sum_{i=1}^{S} p_{spec}\cdot\ln(p_{spec})$	Shannon and Weaver (1949)
Stand Density	Stand Density Index	$\mathrm{SDI} = N_{obs}\cdot\left(\frac{25}{qmd}\right)^{-1.605}$	Reineke (1933)
	Crown Competition factor	$\mathrm{CCF} = \frac{1}{A}\cdot\sum_{i=1}^{z}\left(\frac{cd^2\cdot\pi}{4}\right)\cdot n_{rep}\cdot 100$	Krajicek et al. (1961)
Stand differentiation	Diameter Differentiation Index	$\mathrm{Diff} = 1 - \frac{1}{N_{Obs}}\cdot\sum_{i=1}^{z} n_{rep_i}\cdot\frac{min(dbh_i,dbh_{nextneighbour})}{max(dbh_i,dbh_{nextneighbour})}$	Füldner (1995)

The indicator set considers horizontal distribution, tree species diversity, Stand Density and stand differentiation. To assess the Crown Competition factor, *cd* was computed following Hasenauer (1997).

n_{rep}: Represented stem number in angle count sample

r_a: Distance to nearest tree

N_{obs}: Stem number per hectare

qmd: Quadratic mean diameter

z: Number of trees in an angle count sample

S: Number of tree species

p_{spec}: Proportion of the *i*th tree species

cd: Crown diameter of open grown tree

a_o, a_1: Tree species specific coefficients

dbh: Diameter at breast height

A: Area

Table 2 Scientific evidence for a comprehensive indicator set in a surrogate approach

Structural diversity indicators	Scientific evidence
Clark & Evans-Index	Greater *structural spatial diversity* increases resource partitioning among species (Kohyama 1993; Yachi and Loreau 2007; Álvarez-Yépiz et al. 2017; Atkins et al. 2018) The *variation of tree spacing* provides an indication of the size and distribution of gaps (Neumann and Starlinger 2001) and thus indirectly on processes such as mortality, ingrowth, and competition (Svensson and Jeglum 2001)
Shannon Index	*Tree species abundance* can be used as a proxy for habitat quality or biotope trees (Heym et al. 2021) and related microhabitats (Larrieu et al. 2014) or habitat types (Kovac et al. 2020); e.g., saproxylic beetles, bryophytes, lichens, fungi, and arthropods (Uliczka and Angelstam 1999; Brändle and Brandl 2001; Berglund et al. 2009; Ulyshen 2011) There is high scientific evidence for a positive relation between *tree species diversity* and the number of bird (Baguette et al. 1994; Fisher and Goldney 1998), ground beetle (Fahy and Gormally 1998; Davis et al. 2000; Magura et al. 2000), arthropod (Chey et al. 1997) and ground vegetation species (Fahy and Gormally 1998; Humphrey et al. 2002) *Tree species richness* is a proxy for the number of niche spaces filled by different tree species (Turnbull et al. 2016)
Stand Density Index	The *SDI* can be used as a proxy for spatial distribution of resource availability in biodiversity assessments (Heym et al. 2021) *Gap fraction* indicates the availability of open niche space (McElhinny et al. 2005; LaRue et al. 2019) Compared with CCF, *SDI* is also applicable in mixed forest stands & pure European beech stands (Sdino 1996)
Crown Competition factor	Greater *overlap of crowns* indicates a greater use of niche space for light in the canopy (Williams et al. 2017; Zheng et al. 2015) and can therefore be an indirect measure of ecological niche space (LaRue et al. 2019) In a meta-study, *tree canopy cover* could be related negatively to spider species richness in Europe (Gao et al. 2015) Contrary to the SDI, the *CCF* can delivers reliable outcomes in uneven aged stands (Sterba 1987)
Diameter Differentiation Index	*Variation of tree dimension* can be used as a proxy for habitat quality or biotope trees (Heym et al. 2021) and related macro- and microhabitats (Larrieu et al. 2014), e.g., saproxylic beetles, and lichens (Berglund et al. 2009; Uliczka and Angelstam 1999) *Large tree diameters* indicate high potential for tree related habitats (Hilmo et al. 2009; Nascimbene et al. 2008)

in the BR core areas, as well as forest age classes, tree species browsed, fraying & bark peeling effects, and tree structural foursome were not considered.

Machine learning approach

Random forest models

Random forest models are composed from regression trees and are trained to predict the values of five structural diversity indicators. We are using the statistical language *R* (R Core Team 2020) with the packages *randomForest* (Breiman 2001, 2002), *VSURF* (Geneuer

Table 3 Explanatory variables tested

Category	Explanatory variables
Age and density	Stem density (N), stem basal area (BA), standing stock volume (V), and quadratic mean diameter (qmd)
Vertical structure	Dominant tree species (dom spec), coarse woody debris > 5 cm dm (cwd > 5 cm), coarse woody debris > 25 cm dm (cwd > 25 cm), standing dead wood (sdw), and natural regeneration (regen)
Forest site	Mesorelief (MesoR), microrelief (MicroR), aspect (asp), and altitude (alt)
Game impact	Percentage of regeneration with browsing damage (bd) and amount of young trees without browsing damage (wbd)
Soil and bedrock	Flysch- or limestone forests (flysch), bedrock (bedrock), soil type (soil t), soil moisture (soil m) and humus type (humus)

et al. 2015) and *randomForest explainer* (Ishwaran et al. 2010). In total, 15 random forest models are trained, three for each diversity indicator.

The first random forest models consider 15 explanatory variables (Table 2) of the categories 1–4 (i.e., age & density, vertical structure, forest site, and game impact) per diversity indicator. These models are trained based on data of 1555 permanent sample plots. The second random forest models consider 20 explanatory variables (Table 2) of the categories 1–5 per diversity indicator. These models are trained based on 1555 permanent sample plots. The third random forest models characterize the interrelation between the five structural diversity indicators within the comprehensive indicator set. These models are trained based on data of 422 permanent sample plots including soil monitoring information.

Every random forest is composed of 500 regression trees. For every regression tree, a training set is drawn using bootstrap aggregating (bagging). The decision tree is built by rule-based splitting of the bagging sample into subsets, maximizing the variance between the subsets (Venables and Ripley 2002). At each split in the learning process, a random subset of explanatory variables is used (Ho 1998). The splitting process is repeated recursively on each derived subset, until (i) the subset has identical values with the target variable or (ii) the splitting does no longer add value to the prediction (Quinlan 1986). The mean value of the target variable within a final subset (leaf of a decision tree) is used as the conditional prediction of the target variable for a corresponding combination of explanatory variables (Venables and Ripley 2002).

Variable importance

The importance of every explanatory variable j is assessed by two measures, the percentual increase of the mean squared error (Geneuer et al. 2015; Zhu et al. 2015) and the average minimal depth (Ishwaran et al. 2010). To compute the mean squared error (%IncMSE), the out-of-bag error for every variable j is recorded during the fitting process and averaged over the random forest. Then, the estimated values of j are randomly permutated in the out-of-bag data and dropped down every fitted tree. A higher mean squared error (%IncMSE) indicates higher variable importance and higher explanatory power of the variable. Slightly negatively %IncMSE values may arise in case the mean squared error of the original predictor variable exceeds %IncMSE of permuted values.

To compute the average minimal depth (AvgMinDepth), the level on which variable *j* is used on average to split the decision tree for the first time is assessed. Averaging MinDepth over 500 decision trees yields the average minimal depth (AvgMinDepth) in our case study. Lower AvgMinDepth values indicate higher variable importance and higher explanatory power of the variable.

Variable selection

A two-step variable selection procedure implemented in the R package *VSURF* (Geneuer et al. 2015) is used. VSURF strengthens the models by preselecting a subset of explanatory variables with sufficient explanatory power and removing variables with little or no explanatory power in advance. For details, please see Geneuer et al. (2015).

Results

Levels of structural diversity in the unmanaged core areas

In line with Bitterlich (1984), Lappi and Bailey (1987), Sterba (2008) we aggregated indicator scores on the core area level (Table 4) which is particularly important using angle count method data (Storch et al. 2018).

Indicator–indicandum relationships of forest structural diversity indicators in lowland, European beech forests

Variable importance of explanatory variables is measured by two metrics, %IncMSE and AvgMinDepth. In the text, variables are ordered by the %IncMSE values because differences are more pronounced with this indication. Figures 2, 3, 4, 5, and 6 additionally display AvgMinDepth to gain insights in variable importance distribution among the 500 decision trees. For stand variable abbreviations, please refer to Table 3.

The Clark & Evans-Index (C & E)

In the *first random forest model*, variables indicated best are 'stem density' (%IncMSE = 37.05; AvgMinDepth = 1.73), 'quadratic mean diameter' (%IncMSE = 29.32; AvgMinDepth = 1.92) and 'standing stock volume' (%IncMSE = 28.06; AvgMinDepth = 1.90). In the *second random forest model*, 'stock volume' (%IncMSE = 22.94; AvgMinDepth = 1.42), 'stem basal area' (%IncMSE = 19.33; AvgMinDepth = 2.26), and 'stem density' (%IncMSE = 16.77; AvgMinDepth = 2.28) prove to be most relevant. All variables predicted well by C & E belong to the 'age & density' category. The *third random forest model* detects intercorrelation with the Crown Competition factor (%IncMSE = 13.12; AvgMinDepth = 1.19) and Stand Density Index (%IncMSE = 11.15; AvgMinDepth = 1.4).

The Shannon Index (H′)

In the *first random forest model,* H´ predicts 'dominant tree species' (%IncMSE = 50.45; AvgMinDepth = 1.36) best which belongs to variable category 'vertical structure'

Table 4 Structural diversity levels in the core areas

Nr. core area	Name	Clark & Evans Index			Shannon Index			Stand Density Index			Crown Competition factor			Diameter Differentiation		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
1	Altenberg	0.97	0.27	1.66	0.47	0	1.39	570.6	185.1	1180.1	20.46	1.07	53.47	0.51	0.04	0.98
2	Anninger Tieftal	0.80	0.18	1.64	0.78	0	1.28	696.7	117.1	1433.5	10.02	0.94	34.55	0.48	0.07	0.97
3	Baunzen	1.02	0.15	1.65	0.48	0	1.49	638.1	84.2	1548.2	9.42	0.08	23.27	0.51	0.05	0.98
4	Deutschwald I	1.18	0.55	1.95	0.31	0	1.33	748.8	333.6	1245.0	35.78	10.88	85.59	0.48	0.04	0.98
5	Deutschwald II	1.19	0.84	1.69	0.62	0	1.24	702.3	267.7	1256.8	88.78	22.94	137.76	0.53	0.07	0.98
6	Dombachgraben	0.99	0.49	1.63	0.27	0	1.09	678.1	242.1	1197.0	22.37	0.69	66.28	0.51	0.05	0.97
7	Dorotheerwald	1.04	0.52	2.42	0.63	0	1.27	699.8	216.6	1408.7	10.00	0.25	31.78	0.53	0.10	0.98
8	Festenberg	1.06	0.38	2.29	0.45	0	1.31	586.0	204.9	1088.8	21.19	0.44	70.47	0.32	0.08	0.81
9	Finsterer Gang	1.15	0.68	1.79	0.79	0	1.68	671.4	356.4	1179.3	18.58	3.78	43.91	0.29	0.05	0.71
10	Gießhübl-Eichberg	1.00	0.27	2.15	0.52	0	1.41	521.6	58.8	1063.8	12.65	0.14	29.61	0.71	0.11	0.98
11	Hainbach	1.06	0.44	2.02	0.41	0	1.50	666.3	65.2	1352.0	7.83	0.16	26.67	0.65	0.07	0.98
12	Helenental	1.01	0.14	2.25	0.61	0	1.61	599.4	61.5	1309.0	6.66	0.46	17.19	0.66	0.19	0.97
13	Hengstlberg	1.15	0.59	2.55	0.24	0	1.28	653.4	90.8	1022.5	13.64	0.15	32.77	0.60	0.09	0.98
14	Hirschenstein	1.03	0.38	1.62	0.67	0	1.43	703.1	87.3	1371.0	7.48	0.19	18.04	0.79	0.21	0.98
15	Hoher Lindkogel	0.98	0.14	2.54	0.40	0	1.31	645.7	76.8	1761.0	2.08	0.02	10.62	0.64	0.01	0.98
16	Höherberg	1.05	0.22	1.64	0.67	0	1.44	736.5	291.7	1689.4	15.09	1.91	36.88	0.29	0.10	0.57
17	Hollergraben	1.15	0.50	2.27	0.65	0	1.32	711.0	365.6	1120.2	122.72	37.52	295.81	0.75	0.14	0.98
18	Johannser Kogl I	0.93	0.54	1.56	0.71	0	1.35	475.7	196.2	1009.0	161.79	11.00	527.82	0.63	0.06	0.98
19	Johannserkogl II	0.99	0.43	1.97	0.65	0	1.15	553.9	146.5	1191.7	151.49	3.94	438.27	0.38	0.11	0.78
20 & 21	Kiental I/II	1.05	0.45	1.90	0.51	0	1.39	683.5	187.6	1337.0	13.61	1.77	41.35	0.33	0.08	0.65
22	Kolbeterberg	1.09	0.55	1.92	0.32	0	1.08	541.0	117.0	1228.3	36.80	0.38	144.63	0.44	0.08	0.88
23	Latisberg	1.28	0.68	2.80	0.49	0	1.09	596.7	185.7	901.5	58.81	1.42	166.34	0.55	0.07	0.98
24	Leopoldsberg I	0.95	0.25	1.73	0.57	0	1.56	478.0	157.4	1235.5	15.06	1.15	45.69	0.83	0.58	0.96
25	Leopoldsberg II	0.96	0.47	1.35	0.60	0	1.56	626.6	50.3	1372.8	50.62	7.43	116.96	0.86	0.34	0.98
26	Mauerbach	1.10	0.46	1.72	0.22	0	1.31	657.0	61.0	1078.1	16.91	0.31	32.34	0.44	0.09	0.99

Table 4 (continued)

Nr. core area	Name	Clark & Evans Index			Shannon Index			Stand Density Index			Crown Competition factor			Diameter Differentiation		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
27	Mitterschöpfl	1.08	0.49	1.57	0.35	0	1.28	667.6	112.7	1793.1	9.47	0.09	25.28	0.54	0.08	0.98
28	Moosgraben	1.19	0.50	2.42	0.37	0	1.38	617.4	289.9	999.9	77.88	2.22	200.76	0.80	0.32	0.99
29	Pfaffenberg	1.10	0.43	1.63	0.64	0	1.14	801.5	367.3	1195.5	82.27	17.70	219.36	0.47	0.09	0.87
30	Rauchbuchberg	0.97	0.28	1.74	0.55	0	1.45	567.1	69.8	985.8	53.64	0.35	153.91	0.50	0.06	0.98
31	Sattel	1.03	0.38	1.54	0.43	0	1.24	582.4	97.2	1525.9	21.17	0.28	52.38	0.28	0.00	0.66
32	Schwarzlacken	1.08	0.48	1.63	0.53	0	1.61	678.5	336.5	1390.4	14.50	0.66	32.43	0.36	0.06	0.75
33	Troppberg	1.10	0.20	1.89	0.37	0	1.63	682.6	65.8	1505.4	8.06	0.16	26.29	0.43	0.12	0.98
34	Übelaugraben	1.03	0.51	1.79	0.04	0	0.96	824.1	67.7	1380.1	226.82	14.92	451.33	0.65	0.10	0.98
35	Waldandacht	1.11	0.36	2.09	0.43	0	0.98	747.9	90.8	1350.9	55.02	1.08	159.99	0.56	0.11	0.98
36	Wassergspreng	1.11	0.35	2.17	0.65	0	1.37	644.0	63.0	1558.5	29.19	1.03	84.21	0.26	0.05	0.77
37	Weinberg	1.06	0.20	2.05	0.72	0	1.59	606.4	79.0	994.9	40.49	1.74	92.00	0.25	0.14	0.51
	BR Vienna Woods	1.05	0.14	2.80	0.48	0	1.68	645.3	50.3	1793.1	34.64	0.02	527.82	0.52	0.00	0.99

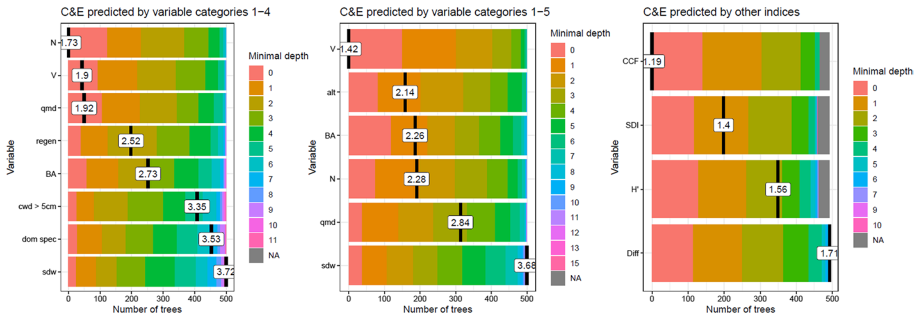

Fig. 2 Indicator–indicandum relationship of the Clark & Evans Index. Minimum depth plots are created by applying the 'R random Forest explainer' package for the *Clark & Evans-Index (C & E)*. The different colors indicate the distribution of the variables' minimal depth (MinDepth) over the 500 decision trees of a random forest. The average minimal depth (AvgMinDepth) of the variables is denoted by the numbers in the white boxes, please note the different scaling. The first random forest model (left panel) considers 1555 permanent sample plots and fifteen explanatory variables, the second random forest model (center panel) considers 422 sample plots and 20 explanatory variables. The third random forest (right panel) considers statistic intercorrelation between the indicators on 1555 permanent sample plots with four explanatory variables

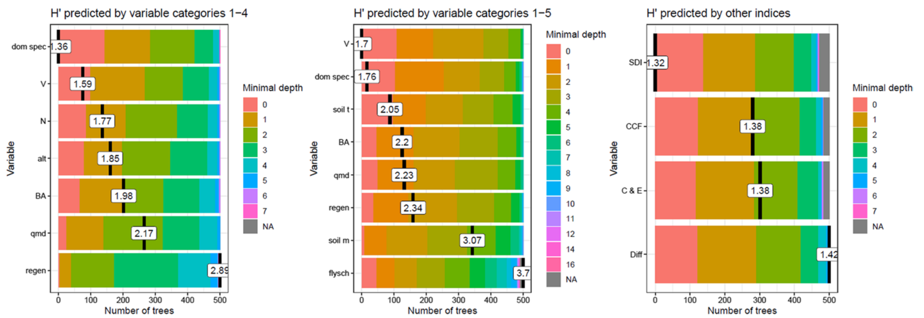

Fig. 3 Indicator–indicandum relationship of the Shannon Index. Minimum depth plots are created by applying the 'R random Forest explainer' package for the *Shannon Index (H')*. The different colors indicate the distribution of the variables' minimal depth (MinDepth) over the 500 decision trees of a random forest. The average minimal depth (AvgMinDepth) of the variables is denoted by the numbers in the white boxes, please note the different scaling. The first random forest model (left panel) considers 1555 permanent sample plots and fifteen explanatory variables, the second random forest model (center panel) considers 422 sample plots and 20 explanatory variables. The third random forest (right panel) considers statistic intercorrelation between the indicators on 1555 permanent sample plots with four explanatory variables

(Fig. 3). This variable is followed by 'standing stock volume' (%IncMSE = 29.94; AvgMinDepth = 1.59) and 'stem density' (%IncMSE = 23.44; AvgMinDepth = 1.77) out of the 'age & density' category. The *second random forest model* indicates highest variable importance for 'dominant tree species' (%IncMSE = 25.63; AvgMinDepth = 1.76), 'standing stock volume' (%IncMSE = 14.74; AvgMinDepth = 1.70) and 'soil type' (%IncMSE = 14.44; AvgMinDepth = 2.05). In the *third random forest model*, H′ reveals closest statistical relation to the Stand Density Index (%IncMSE = 22.95; AvgMinDepth = 1.32) and the Crown Competition factor (%IncMSE = 21.49; AvgMinDepth = 1.38) within the indicator set.

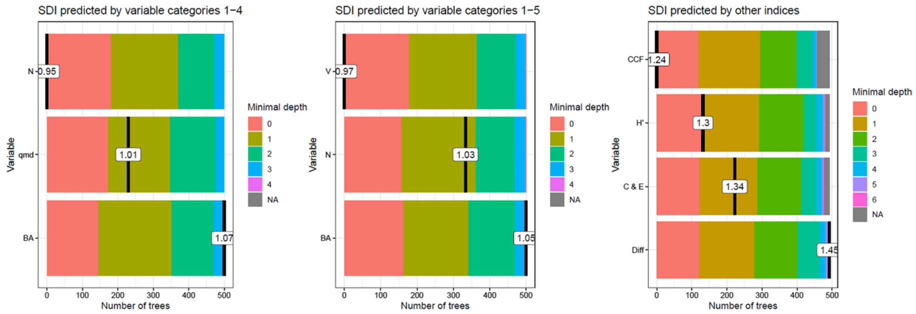

Fig. 4 Indicator–indicandum relationship of the Stand Density Index. Minimum depth plots are created by applying the 'R random Forest explainer' package for the *Stand Density Index (SDI)*. The different colors indicate the distribution of the variables' minimal depth (MinDepth) over the 500 decision trees of a random forest. The average minimal depth (AvgMinDepth) of the variables is denoted by the numbers in the white boxes, please note the different scaling. The first random forest model (left panel) considers 1555 permanent sample plots and fifteen explanatory variables, the second random forest model (center panel) considers 422 sample plots and 20 explanatory variables. The third random forest (right panel) considers statistic intercorrelation between the indicators on 1555 permanent sample plots with four explanatory variables

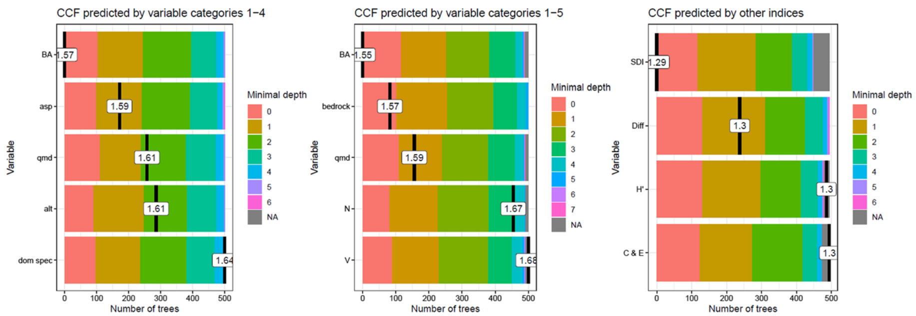

Fig. 5 Indicator–indicandum relationship of the Crown Competition factor. Minimum depth plots are created by applying the 'R random Forest explainer' package for the Crown Competition factor (CCF). The different colors indicate the distribution of the variables' minimal depth (MinDepth) over the 500 decision trees of a random forest. The average minimal depth (AvgMinDepth) of the variables is denoted by the numbers in the white boxes, please note the different scaling. The first random forest model (left panel) considers 1555 permanent sample plots and fifteen explanatory variables, the second random forest model (center panel) considers 422 sample plots and 20 explanatory variables. The third random forest (right panel) considers statistic intercorrelation between the indicators on 1555 permanent sample plots with four explanatory variables

The Stand Density Index (SDI)

All variables indicated well by SDI in the *first random forest model* i.e., 'stem basal area' (%IncMSE = 82.79; AvgMinDepth = 1.07), 'stem density' (%IncMSE = 36.44; AvgMinDepth = 0.95), and 'quadratic mean diameter' (%IncMSE = 31.75%; AvgMinDepth = 1.01) belong to the 'age & density' category. In the *second SDI model*, 'stem density' (%IncMSE = 53.16; AvgMinDepth = 1.03), 'stem basal area' (%IncMSE = 42.02; AvgMinDepth = 1.05), and 'standing stock volume' (%IncMSE = 29.95; AvgMinDepth = 0.97) can be very well predicted. SDI shows closest interrelations to other structural diversity indicators.

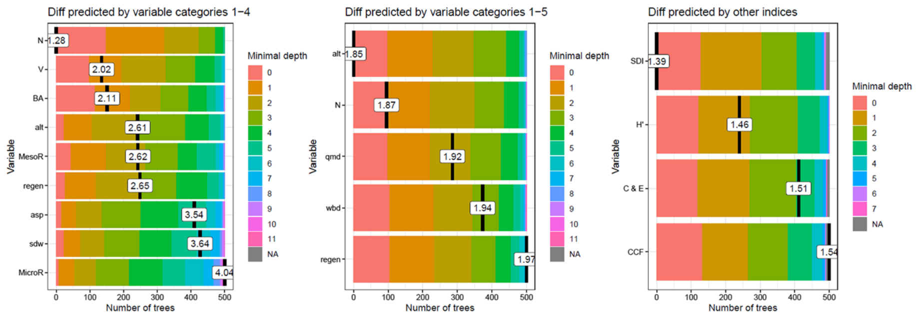

Fig. 6 Indicator–indicandum relationship of the Diameter Differentiation Index. Minimum depth plots are created by applying the 'R random Forest explainer' package for the Diameter Differentiation Index (Diff). The different colors indicate the distribution of the variables' minimal depth (MinDepth) over the 500 decision trees of a random forest. The average minimal depth (AvgMinDepth) of the variables is denoted by the numbers in the white boxes, please note the different scaling. The first random forest model (left panel) considers 1555 permanent sample plots and fifteen explanatory variables, the second random forest model (center panel) considers 422 sample plots and 20 explanatory variables. The third random forest (right panel) considers statistic intercorrelation between the indicators on 1555 permanent sample plots with four explanatory variables

Adding any of the remaining indicators to the *third random forest model* yields %IncMSE between 13 and 35%.

The Crown Competition factor (CCF)

Variables indicated best in the *first CCF random forest model* are 'stem basal area' (%IncMSE = 62.10; AvgMinDepth = 1.57), 'dominant tree species' (%IncMSE = 61.22; AvgMinDepth = 1.64) and 'quadratic mean diameter' (%IncMSE = 47.69; AvgMinDepth = 1.61). In *the second model*, four variables prove high explanatory power, namely 'stem density' (%IncMSE = 28.99; AvgMinDepth = 1.67), 'stem basal area' (%IncMSE = 27.35; AvgMinDepth = 1.55), 'quadratic mean diameter' (%IncMSE = 27.04; AvgMinDepth = 1.59), and 'standing stock volume' (%IncMSE = 25.77; AvgMinDepth = 1.68). The *third random forest model* detects closest intercorrelation between CCF and SDI (%IncMSE = 33.18; AvgMinDepth = 1.29).

The Diameter Differentiation Index (diff)

Three explanatory variables, all belonging to the category 'age & density', are indicated best by Diff in the *first random forest model:* 'Stem density' (%IncMSE = 25.46; AvgMinDepth = 1.28), 'standing stock volume' (%IncMSE = 19.53; AvgMinDepth = 2.02), and 'stem basal area' (%IncMSE = 18.39; AvgMinDepth = 2.11). In the *second random forest model*, variables predicted well are 'quadratic mean diameter' (%IncMSE = 16.36; AvgMinDepth = 1.92) and 'stem density' (%IncMSE = 15.43; AvgMinDepth = 1.87). The *third random forest model* displays closest intercorrelation to the SDI (%IncMSE = 15.13; AvgMinDepth = 1.39) and the Crown Competition factor (%IncMSE = 14.87; AvgMinDepth = 1.54).

Indicator–Indicandum relationships of aa comprehensive forest biodiversity indicator set

The variable category neglected by the indicator set are ‘game impact’ and ‘soil & bedrock’ (Fig. 7). Partially reflected are the categories ‘forest site’ and ‘vertical structure’. Variables of the category ‘age & density’ are overrepresented. There are no major differences between first and second model results. However, variable importance decreases on average about − 23% in the second models compared to the first ones which consider a lower number of explanatory variables. Testing fifteen instead of twenty explanatory variables affects the sum of explanatory power between + 9%IncMSE (SDI) and + 33%IncMSE (H′). Using randomForest to gain insight in indicandum–indicator relationships, a pronounced sensitivity to the number of explanatory variables tested could be found.

Explanatory variables indicated best by the indicator set in the first and second models are stem basal area (BA = 154.43%IncMSE), stem density (N = 117.61%IncMSE), standing stock volume (V = 101.34%IncMSE), quadratic mean diameter (qmd = 81.02%IncMSE),

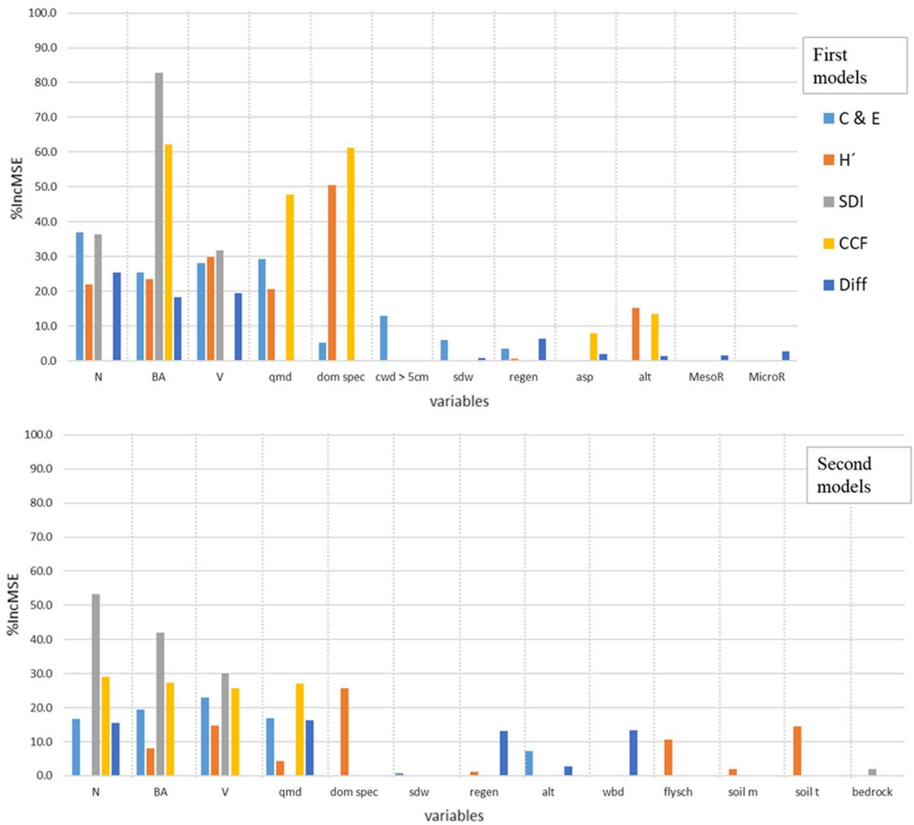

Fig. 7 Indicator–indicandum relationships of a comprehensive indicator set. Overview of the mean squared error (%IncMSE) created with R random Forest to indicate explanatory variable importance for the indicator set, consisting of Clark & Evans-Index (C & E), Shannon Index (H′), Stand Density Index (SDI), Crown Competition factor (CCF) and Diameter Differentiation Index (Diff). Upper panel: First random forest models (1555 sample plots, 15 explanatory variables). Lower panel: Second random forest models (422 sample plots, 20 explanatory variables)

and dominant tree species (dom spec = 71.22). 17 of 20 explanatory variables under study are at least once indicated in the ten models. The three stand variables overall neglected by the indicators set are coarse woody debris < 25 mm MDM (cwd < 25 mm), proportion of regeneration with browsing damage (bd), and humus type (humus).

Intercorrelation within a comprehensive indicator set

The five structural indicators are highly interrelated (Fig. 8). Overall, SDI shows highest statistical relation to other diversity indicators and can also be well predicted by them. Moreover, Adding CCF to a model considerably raises it's explanatory power. Contrary, C & E displays very low statistic relation to other structural diversity indicators. Overall highest correlation can be found between SDI and CCF (33.2–34.8%IncMSE). Strong correlations within indicator sets may arise due to description of the same structural aspect (e.g., Stand Density) and by sharing direct elements (e.g., tree diameter and stem density) in the formula.

Discussion

Model approach

Comparing the use of a machine learning approach (*random Forests*) to gain additional insights into indicator–indicandum relationships and intercorrelation within indicator sets in comparison to e.g., linear regression with forward selection, we see following main advantage for ecological science: (i) no assumptions about linear relationships are needed, (ii) a possible collinearity of variables does not affect model predictions negatively and (iii) stable prediction results in terms of the Out-Of-Bag error. The disadvantages of *random Forests* are that (i) outcomes are more challenging to interpret, (ii) direction of statistic relation is unknown, and (iii) collinearity might affect %IncMSE, are clearly outweighed in our case study. The package *random Forest explainer* proved to be a useful tool to interpret the model outcomes.

Validity of most indicators used is weakly scientifically supported (Gao et al. 2015). A biodiversity indicator is found to be more useful the more precise the correlation with the indicandum is known (Heink and Kowarik 2010). Yet, indicator–indicandum relationships

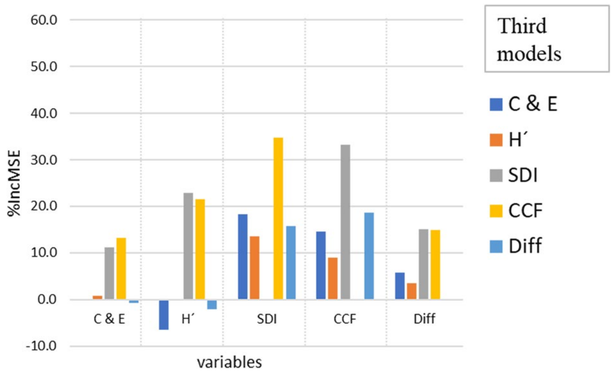

Fig. 8 Overview of intercorrelation within the comprehensive indicators set. Overview of the mean squared error (%IncMSE) created with R random Forest to characterize the interrelation between Clark & Evans-Index (C & E), Shannon Index (H´), Stand Density Index (SDI), Crown Competition factor (CCF) and Diameter Differentiation (Diff)

are poorly understood and tested across habitat and scales (Gao et al. 2015). Our case study shows, how *random Forest* can be applied for the indicator validation urgently needed on large spatial scales (Ferris and Humphrey 1999; Gao et al. 2015) considering intercorrelated data and indicators sets.

Indicator–indicandum relationships

The Clark & Evans-Index (C & E)

Actual C & E levels in the unmanaged core areas of the BR Vienna Woods range between 0.76 (‘Anninger’) and 1.25 (‘Latisberg’). In the core area ‘Anninger’, trees are evenly arranged, while stem distribution in ‘Latisberg’ already evolved towards a more clustered spatial structure. Comparable C&E levels to ‘Latisberg’ were found in a 53-year-old pure European beech stand in Germany (Pommerening 2002). Older stands tend to have lower stem numbers and clumped structure, while young stands are often found to be evenly arranged (Pretzsch 2002; Dieler 2013). Even if mean stand age only differs about 20 years between two core areas, ‘Latisberg’ displays twice the amount of living stock volume and half the number of trees per hectare. In line with Pretzsch (2002) and Dieler (2013), this points towards a more mature successional state of ‘Latisberg’ which is indicated by C & E.

In unmanaged forests, structural complexity, and diversity significantly increase with stand age, denoted by enhanced levels of lying and standing deadwood and natural regeneration (Pretzsch 2002). In line with the findings of Pretzsch (2002) all these variables (cwd, swd, regen) are very well indicated by C & E in our case study. C & E indicates the variable category ‘vertical structure’ very well. Moreover, our results underline a profonde indication of the category ‘age & density’.

C & E was found to indicate horizontal distribution as a proxy for resource partitioning of light use among species (Kohyama 1993; Yachi and Loreau 2007; Álvarez-Yépiz et al. 2017; Atkins et al. 2018), the size and distribution of gaps (Neumann and Starlinger 2001) and processes such as mortality, ingrowth, and competition (Svensson and Jeglum 2001). Therefore, it is highly plausible that the variable indicated best by C& E in our case study is ‘stem density’. C & E shows particularly low statistical relation to other structural diversity indicators.

The Shannon Index (H′)

Shannon Index levels varies between 0.04 (‘Übelaugraben’) and 0.79 (‘Finsterer Gang’). Comparable Index levels were described for pure European beech forest (H′ = 0.09) and oak-beech mixed forest (H′ = 0.62) in Germany (Pommerening 2002). Rare species increase H′ disproportionately, while common species affect it under proportionately (Pretzsch 2002). Overall, the Shannon Index indicates the highest number of variables. Moreover, the category ‘vertical structure’ and the variable ‘dominant tree species’ are predicted best by the Shannon Index. This is supported by scientific literature in which the Shannon Index is expected to indicate tree species abundance and diversity and is considered as a proxy for the number of niche spaces filled by different tree species (Turnbull et al. 2016), habitat quality or biotope trees (Heym et al. 2021), diversity of microhabitats (Larrieu et al. 2014), and habitat types (Kovac et al. 2020) for a variegation of taxonomic groups.

Of the five diversity indicators surveyed in this study, H′ indicates variables of ‘soil & bedrock’ best (s. s., soil type, flysch or limestone Vienna Woods, and soil moisture). These variables are interdependent, have major impact on plant communities, and underline the geological peculiarity of the study area. The distinction between flysch and limestone Vienna Woods has crucial implications for the soil types, their chemical composition and water balance, as well as the diversity of occurring animal and plant species (BFW 2011). In the flysch parts of the Vienna Woods, heavy, nutrient-rich, deep soils have developed. These soils are characterized by advantageous water supply and high specific water storage capacity (Leitgeb et al. 2012). Species diversity monitoring in the BR Vienna Woods detects few vascular plant species in high abundances in those areas (BR Vienna Woods Management 2021a). In the limestone parts of the study area, dry, nutrient-poor, and shallow soils are common (BFW 2011). Specific water storage capacity and water supply of these soils is much lower and promote drought tolerance species (Leitgeb et al. 2012). Species diversity monitoring indicates species-rich herbaceous vegetation in low abundances (BR Vienna Woods Management 2021a) making the model outcomes highly reasonable.

The Stand Density Index (SDI)

Of all indicators, the variables ‘basal area’, ‘living wood volume’, ‘quadratic mean diameter’, and ‘Stand Density’ are indicated best by SDI. The Stand Density Index reflects the lowest number of explanatory variables, all belonging to the category ‘age & density’, in very high accuracy. In our study, pronounced correlations with other indicators, especially with CCF, are found. Besides directly sharing the element ‘Stand Density’ in their formula, CCF and SDI describe the same forest structural aspect. The SDI is a proxy for spatial distribution of resource availability (Heym et al. 2021) and indicates the availability of open niche space (McElhinny et al. 2005; LaRue et al. 2019). Actual SDI levels in the core areas of the BR Vienna Woods range between 524.69 (‘Leopoldsberg I’) and 877.51 (‘Rauchbuchberg’). Our findings line up with Vospernik and Sterba (2016) who demonstrated maximum stand densities stands of tree species in Austria. Pure coniferous and mixed stands show comparably higher Stand Density levels than broadleaved stands.

No correlation between SDI and ‘dominant tree species’ is detected in the case study, even if e.g., tree mortality with increasing Stand Density was found to be strongly tree species dependent (Liang et al. 2007). This indicates that (1) the broadleaved species observed have similar maximum densities in terms of stem numbers and basal areas or (2) species dependent mortality does not yet play a major role in the core areas of the BR Vienna Woods.

Additionally, occurrence of ‘clastic bedrock’, on which nutrient-poor soils establish (NW-FVA 2008), is indicated by SDI. Our study shows how canopy competition in the BR Vienna Woods could be a proxy for soil nutrient supply. These findings are in line with Schmidt et al. (2002) and Podrázský et al. (2014), who proved that soil base supply is the most important factor explaining herbaceous species diversity in temperate beech and Douglas fir forests. Greater overlap of crowns indicates a greater use of niche space for light in the canopy (Williams et al. 2017), and limits light transmission to the ground. In future studies, it would hence be interesting to test if ground vegetation diversity or quantity can be indicated by SDI in European beech dominated forests.

The Crown Competition factor (CCF)

The ranking of the core areas deviates between Stand Density assessment with SDI and CCF. Actual CCF levels in the core areas of the BR Vienna Woods range between 225.60 ('Johannserkogel II') and 471.75 ('Übelaugraben'). CCF can be applied to uneven-aged mixed forests (Sterba 2008). Difficulties with CCF can arise with the assessment of pure stands of the very shade-tolerant and large-crowned European beech, for which Sdino (1996) described maximum CCF levels of > 2000. Variables well indicated by CCF are 'stem basal area', 'quadratic mean diameter', and 'dominant tree species'. The indication of 'dominant tree species' by the CCF is in line with Sdino (1996) and Liang et al. (2007) and may occur due to the species-wise crown diameter being considered in the CCF formula.

Moreover, CCF indicates the variables 'altitude' and 'aspect' well. The Vienna Woods contains both, hall-shaped, low understory beech stands and south-exposed hilltops, where European beech (*Fagus sylvatica*) is already water-limited. On those sites, red pine and oak forest communities with rich understory occur (BR Vienna Woods Management 2021b), making this result highly plausible.

The Diameter Differentiation Index (Diff)

The Diameter Differentiation Index is the only indicator to mirror game impact and an overall high number of variables. Closest intercorrelation of Diff is found with CCF and SDI, both of which Diff shares one element (qmd) in the formula with, respectively. Actual Diameter Differentiation Index levels in the core areas of the BR Vienna Woods range between 0.22 ('Hengstlberg') and 0.40 ('Johannserkogel I'). Diameter heterogeneity in unmanaged stands is created by natural disturbance regimes which are decisive for most forest structural legacies. Natural disturbance regimes of European beech forests contain frequent, small-scale, low intensity as well as rare, large-scale, high intensity disturbance events (Leibundgut 1982; Mayer 1984; Tabaku 1999; Meyer et al. 2003).

Species diversity monitoring in the BR Vienna Woods shows that occurrence probabilities for bat, snail, relict beetle, and old-growth forest bird species increase in the core areas compared to the managed parts. The Diameter Differentiation Index seems to mirror plenty of the crucial habitat structures and quality for those guilds best (e.g., altitude, aspect, micro- and meso-relief, natural regeneration and standing dead wood). Deadwood input often relates with the natural disturbance regimes (Christensen et al. 2005). The outcomes line up with findings of Winter and Möller (2008) who showed that the Diff can be an important indicator of microhabitats in forest stands.

Indicative value of a comprehensive biodiversity indicator set

The variable category 'age & density' is overrepresented by the comprehensive indicator set. Partially reflected are the categories 'forest site' and 'vertical structure'. The categories neglected are 'game impact' and 'soil & bedrock'. Using random Forest to gain new insights in indicandum–indicator relationships, pronounced sensitivity to the number of explanatory variables tested could be found. Variables reflected best by the indicator set are 'stem basal area', 'stem density', 'standing stock volume', and 'quadratic mean diameter'. Contrary, stand characteristics like 'coarse woody debris > 25 MDM', 'tree browsing', and 'humus type' are neglected in all models. Scientifically, there is broad consensus for the relevance of humus type (e.g., Schäfer and Schauermann 1990; Hooper et al. 2000; Ponge 2003; Salmon et al. 2006, 2008), tree browsing (e.g., Gill 1992; Pastor et al. 1997; Reimoser et al. 2003) and large coarse woody debris (Kappes & Topp 2004; Müller et al. 2007; Rondeux and Sanchez 2010; Brin et al. 2011; Lassauce et al. 2011) for forest biodiversity.

In line with LaRue et al. (2019), our study shows that aspects of forest structure indeed are intercorrelated and neither ecologically nor statistically independent. Furthermore, we agree with these authors that structural niche space or ecosystem structure and function cannot be understood by one metric. Indicators which measure either more or less than they are supposed to, i.e., construct-irrelevant variance or construct underrepresentation may bias the qualitative connection between evidence and interpretation (Heink and Kowarik 2010).

Due to unavailable indicator values (e.g., bark diversity, hollow trees, forest communities, litter dry weight, litter decomposition, perennial species richness, tree age, and undisturbed reference areas) or different scales it was not possible to compare our indicator set with the performance of other aggregated biodiversity indicators (Parkes et al. 2003; McElhinny et al. 2006; Geburek et al. 2010; Storch et al. 2018; Heym et al. 2021). However, there is partial agreement in choice of elements of biodiversity studied in McElhinny et al. (2006) and Storch et al. (2018) like quadratic mean diameter, natural regeneration, standing and lying deadwood, stem basal area. Compared to Heym et al. (2021) partly identical structural diversity indicators are chosen (e.g., Shannon Index, SDI).

Handling knowledge gaps in biodiversity monitoring by machine learning approaches has already been explored in permanent grassland and freshwater ecosystems (Gallardo et al. 2011; Plantureux et al. 2011). In line with these authors, our case study underlines the large potential of machine learning for testing indicative value of single indicators and comprehensive forest biodiversity indicator sets. Moreover, machine learning could advance biodiversity indicator choice.

Summary and conclusion

In this publication, a machine learning approach to provide novel insights in indicator–indicandum relationships of biodiversity indicators and comprehensive indicator sets is presented. The indicators tested are parameters of forest spatial and structural heterogeneity. We surveyed a comprehensive indicator set of Clark & Evans-, Shannon, Stand Density, Diameter Differentiation Index, and Crown Competition factor with randomForest and examine their indicative value for twenty explanatory stand variables.

Biodiversity indicators are sometimes criticized for displaying poor indicator–indicandum relationships (Ferris and Humphrey 1999; Margules et al. 2002; Duelli and Obrist 2003; Gao et al. 2015). Machine learning proves to be a useful tool to overcome these knowledge gaps and provides additional insights in indicator–indicandum relationships. This scientific work deepens understanding of statistic properties of forest-inventory based biodiversity indicators and comprehensive indicator sets.

Examining 37 unmanaged core areas in the Vienna Woods, following scientific questions are answered: Which levels of structural diversity can be found in the unmanaged core areas of the Biosphere Reserve Vienna Woods? (2) Which stand characteristics are indicated by single structural diversity indicators? (3) Which stand characteristics are indicated or neglected by a comprehensive indicator set? (4) How strong is the intercorrelation in an indicator set?

Indicator choice is the most crucial step in biodiversity assessments. In our study, the Shannon Index is found to be most useful to indicate the variable category ‘soil & bedrock’ and ‘vertical structure’. Variables of ‘age & density’ are best considered using the Stand Density Index which indicates a low number of stand variables in very high accuracy. CCF indicates the variables of ‘forest site’ best and altogether displays closest relation to all variables studied. The Diameter Differentiation Index is the only indicator to mirror ‘game impact’ and might reflect natural disturbance regimes well. Overall, the Shannon Index indicates highest, the Stand Density Index lowest number of forest stand characteristics.

Strong correlations between indicators may arise due to indication of the same forest structural aspect in indicator sets and/or by sharing direct elements in the formula. To rise reliability of biodiversity assessments, both should most possibly be avoided. Some stand characteristics (e.g., variable category ‘age & density’) relevant to biodiversity are indicated disproportionally in the comprehensive indicator set, while other important ones (e.g., ‘coarse woody debris < 25 MDM’, ‘tree browsing’, and ‘humus type’) are neglected.

More ecological studies are needed to explore indicator–indicandum relationships in detail. Machine learning as integral part of artificial intelligence may be a novel, effective and entire objective way to gain new insights into indicator–indicandum relationships on variable scales. The prediction outcome is decisively impacted by type and number of explanatory variables tested. The smaller the number of input variables, the more parsimonious is the model. Preselecting variables with regression algorithms is highly recommended. Random Forest models assumes interval scaled variables. Therefore, the impact of interval-scaled, common features on biodiversity can effectively be evaluated with machine learning. Nonetheless, relevance of qualitative variables and rare events may be underestimated. The methodology described in this study might be more suitable to review quantitative (measurable) than qualitative (observed) variables.

Our goal was to contribute to the use of inventory-based structural diversity indicators in forests by precising indicator–indicandum relationships through machine learning. This case study shows, how random forest models can be applied for the indicator validation on large spatial scales, considering intercorrelated data and comprehensive sets of structural diversity indicators. It might be a useful tool to create novel biodiversity indicator sets. Our findings support the great potential of random Forest in the context of forest biodiversity assessments and indicator choice.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10531-022-02536-0.

Author contributions S.E. curated data and wrote the original manuscript draft. S.V. and T.R. supported manuscript preparation. All authors contributed to data analysis and preparation of figures and tables.

Funding Open access funding provided by University of Natural Resources and Life Sciences Vienna (BOKU). Open access funding provided by University of Natural Resources and Life Sciences Vienna (BOKU). The authors declare that no other funds, grants, or other support were received during the preparation of this manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Álvarez-Yépiz JC, Búrquez A, Martínez-Yrízar A, Teece M, Yépez EA, Dovciak M (2017) Resource partitioning by evergreen and deciduous species in a tropical dry forest. Oecologia 183:607–618

Atkins JW, Fahey RT, Hardiman BH, Gough CM (2018) Forest canopy structural complexity and light absorption relationships at the subcontinental scale. JGR Biogeosci 123:1387–1405. https://doi.org/10.1002/2017JG004256

Baguette M, Deceuninck B, Muller Y (1994) Effects of spruce afforestation on bird community dynamics in a native broadleaved forest area. Acta Oecol 15:275–288

Balvanera P, Pfisterer AB, Buchmann N et al (2006) Quantifying the evidence for biodiversity effects on ecosystem functioning and services. Ecol Lett 9:1146–1156. https://doi.org/10.1111/j.1461-0248.2006.00963.x

Begon M, Harper JL, Townsend CR (1996) Ecology: individuals, populations and communities. Blackwell, Brookline Village

Berglund H, O'Hara RB, Jonsson BG (2009) Quantifying habitat requirements of tree-living species in fragmented boreal forests with Bayesian methods. Conserv. Biol. 23:1127–1137. https://doi.org/10.1111/j.1523-1739.2009.01209.x

BFW (2011) Soil Monitoring in the core areas of the Biosphere Reserve Vienna Woods: final report. Austrian Federal Research Centre for forests, Vienna **(In German)**

BR Vienna Woods Management (2011) Core area monitoring 2011. Biosphere Reserve Vienna Woods Management. https://www.bpww.at/de/aktivitaeten/basis-monitoring-in-den-kernzonen-des-biosphaerenpark-wienerwald. Accessed 20 Jan 2021.

BR Vienna Woods Management (2021a) Biodiversity monitoring in the Biosphere Reserve Vienna Woods. Biosphere Reserve Vienna Woods Management. https://www.bpww.at/de/aktivitaeten/biodiversitaetsmonitoring-und-beweissicherung-in-den-kernzonen-des-biosphaerenpark. Accessed 20 Jan 2021a

BR Vienna Woods Management (2021b) Core area monitoring in the Biosphere Reserve Vienna Woods. Biosphere Reserve Vienna Woods Management. https://www.bpww.at/de/aktivitaeten/basis-monitoring-in-den-kernzonen-des-biosphaerenpark-wienerwald. Accessed 20 Jan 2021b

Brändle M, Brandl R (2001) Species richness of insects and mites on trees: expanding Southwood. J Anim Ecol 70:41–504

Brin A, Bouget C, Brustel H, Jactel H (2011) Diameter of downed woody debris does matter for saproxylic beetle assemblages in temperate oak and pine forests. J Insect Conserv Divers 15:653–669. https://doi.org/10.1007/s10841-010-9364-5

Bitterlich W (1984) The relascope idea: relative measurements in forestry. Commonwealth Agricultural Bureaux, Farnham Royal

Breiman NL (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Breiman NL (2002) Manual on setting up, using, and understanding random forests V3.1. https://www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf. Accessed 20 Jan 2021.

CBD (2014) Global biodiversity outlook 4. Secretariat of the Convention on Biological Diversity, Montréal

CBD (2018) Update on progress in revising/updating and implementing national biodiversity strategies and action plans, including national targets. Secretariat of the Convention on Biological Diversity, Montréal

Chey VK, Holloway JD, Speight MR (1997) Diversity of moths in forest plantations and natural forests in Sabah. Bull Entomol Res 87:371–385. https://doi.org/10.1017/s000748530003738x

Chirici G, Winter S, McRoberts RE (2011) National forest inventories: contributions to forest biodiversity assessments. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0482-4

Christensen M, Hahn K, Mountford EP, Ódor P, Standovár T, Rozenbergar D, Diaci J et al (2005) Dead wood in European beech (*Fagus sylvatica*) forest reserves. For Ecol Manage 210:267–282. https://doi.org/10.1016/j.foreco.2005.02.032

Ceballos G, Garcia A, Ehrlich PR (2010) The sixth extinction crisis: Loss of animal populations and species. J Cosmol 8:1821–1831

Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palme TM (2015) Accelerated modern human–induced species losses: Entering the sixth mass extinction. Sci Adv 1(5):e1400253–e1400253. https://doi.org/10.1126/sciadv.1400253
Clark PJ, Evans FC (1954) Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology 35:445–453. https://doi.org/10.2307/1931034
Corona P, Köhl M, Marchetti M (2003) Advances in forest inventory for sustainable forest management and biodiversity monitoring. Springer, Dordrecht
Corona P, Chirici G, McRoberts RE, Winter S, Barbati A (2011) Contribution of large-scale forest inventories to biodiversity assessment and monitoring. For Ecol Manage 262:2061–2069. https://doi.org/10.1016/j.foreco.2011.08.044
Davis AJ, Huijbregts H, Krikken J (2000) The role of local and regional processes in shaping dung beetle communities in tropical forest plantations in Borneo. Glob Ecol Biogeogr Lett 9:281–292
Dieler J (2013) Biodiversity and Forest Management: Effects on species diversity, structural diversity, and productivity. DVFFA Beiträge zur Jahrestagung **(In German)**
Duelli P, Obrist MK (2003) Biodiversity indicators: The choice of values and measures. Agr Ecosyst Environ 98:87–98. https://doi.org/10.1016/S0167-8809(03)00072-0
EHYD (2021) The electronic hydrographic yearbook of Austria. Austrian Federal Ministry of Agriculture, Regions and Tourism. www.ehyd.gv.at. Accessed 20 Jan 2021.
Ette JS, Geburek T (2021) Why European biodiversity reporting is not reliable. Ambio 50:929–941. https://doi.org/10.1007/s13280-020-01415-8
Fahy O, Gormally M (1998) A comparison of plant and carabid beetle communities in Irish oak woodland with a nearby conifer plantation and clear-felled site. For Ecol Manage 110:263–273. https://doi.org/10.1016/s0378-1127(98)00285-0
FAO (2020) Global forest resources assessment 2020: main report. Food and Agriculture Organization of the United Nations, Rome
Ferris R, Humphrey JW (1999) A review of potential biodiversity indicators for application in British forests. Forestry 72:313–328. https://doi.org/10.1093/forestry/72.4.313
Fisher AM, Goldney DC (1998) Native forest fragments as critical bird habitat in a softwood forest landscape. Aust for 61:287–295. https://doi.org/10.1080/00049158.1998.10674753
Füldner K (1995) Structural characterization of Beech-Hardwood-mixed forests. Dissertation, University of Göttingen.
Gabler K, Schadauer K (2007) Some approaches and designs of sample-based national forest inventories. Aust J for Sci 124:105–133
Gallardo B, Gascónb S, Quintana X, Comín FA (2011) How to choose a biodiversity indicator: redundancy and complementarity of biodiversity metrics in a freshwater ecosystem. Ecol Ind 11:1177–1184. https://doi.org/10.1016/j.ecolind.2010.12.019
Gamfeldt L, Snäll T, Bagchi R, Jonsson M, Gustafsson L, Kjellánder P, Ruíz-Jean MC et al (2013) Higher levels of multiple ecosystem services are found in forests with more tree species. Nat Commun 4:1340–1340. https://doi.org/10.1038/ncomms2328
Gao T, Nielsen AB, Hedblom M (2015) Reviewing the strength of evidence of biodiversity indicators for forest ecosystems in Europe. Ecol Ind 57:420–434. https://doi.org/10.1016/j.ecolind.2015.05.028
Geburek T, Milasowszky N, Frank G, Konrad H, Schadauer K (2010) The Austrian Forest Biodiversity Index: all in one. Ecol Ind 10:753–761. https://doi.org/10.1016/j.ecolind.2009.10.003
Geneuer R, Poggi JM, Tuleau-Malot C (2015) VSURF: an R package for variable selection using random forests. R J. https://doi.org/10.32614/RJ-2015-018
Gill RMA (1992) A review of damage by mammals on north temperate forests III: impact on trees and forests. Forestry 65:363–388. https://doi.org/10.1093/forestry/65.4.363-a
Hasenauer H (1997) Dimensional relationships of open-grown trees in Austria. For Ecol Manage 96:197–206. https://doi.org/10.1016/S0378-1127(97)00057-1
Heink U, Kowarik I (2010) What criteria should be used to select biodiversity indicators? Biodivers Conserv 19:3769–3797. https://doi.org/10.1007/s10531-010-9926-6
Heym M, Uhl E, Moshammer R, Dieler J, Stimm K, Pretzsch H (2021) Utilising forest inventory data for biodiversity assessment. Ecol Ind 121:1–11. https://doi.org/10.1016/j.ecolind.2020.107196
Hilmo O, Holien H, Hytteborn H, Ely-Aastrup H (2009) Richness of epiphytic lichens in differently aged Picea abies plantations situated in the oceanic region of Central Norway. Lichenologist 41:97–108. https://doi.org/10.1017/S0024282909007865
Ho TK (1998) The random subspace method for constructing decision forests. IEEE Trans Pattern Anal Mach Intell 20:832–844. https://doi.org/10.1109/34.709601
Hooper DU, Bignell DE, Brown VK, Brusaard L, Dangerfield JM, Wall DH, Wardle DA et al (2000) Interactions between aboveground biodiversity in terrestrial ecosystems: patterns, mechanisms, and

feedbacks. Bioscience 50:1049–1061. https://doi.org/10.1641/0006-3568(2000)050[1049:IBAABB]2.0.CO;2
Hooper DU, Chapin FS, Ewel JJ, Hector A, Inchausti P, Lavorel S, Lawton JH et al (2005) Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. Ecol Monogr 75:3–35. https://doi.org/10.1890/04-0922
Humphrey JW, Ferris R, Jukes MR, Peace AJ (2002) The potential contribution of conifers plantations to the UK Biodiversity Action Plan. Bot J Scotl 54:49–62. https://doi.org/10.1080/03746600208685028
IPBES (2019) Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES Secretariat, Bonn
Isbell F, Calcagno V, Hector A et al (2011) High plant diversity is needed to maintain ecosystem services. Nature 477:199–202. https://doi.org/10.1038/nature10282
Ishwaran H, Kogalur UB, Gorodeski EZ, Minn AJ, Lauer MS (2010) High-dimensional variable selection for survival data. J Am Stat Assoc 105:205–217. https://doi.org/10.1198/jasa.2009.tm08622
Jones JP, Collen B, Atkinson G, Baxter PW, Bubb P, Illian JB, Katzner TE et al (2011) The why, what, and how of global biodiversity indicators beyond the 2010 target. Conserv Biol 25:450–457. https://doi.org/10.1111/j.1523-1739.2010.01605
Kappes H, Topp W (2004) Emergence of Coleoptera from deadwood in a managed broadleaved forest in Central Europe. Biodivers Conserv 13:1905–1924. https://doi.org/10.1023/B:BIOC.0000035873.56001.7d
Katzner T, Milner-Gulland EJ, Bragin E (2007) Using modelling to improve monitoring of structured populations: are we collecting the right data? Conserv Biol 21:241–252. https://doi.org/10.1111/j.1523-1739.2006.00561.x
Keenan RJ, Reams GA, Achard F, Freitas JV, Grainger A, Lindquist E (2015) Dynamics of global forest area: results from the FAO global forest resources assessment 2015. For Ecol Manage 352:9–20. https://doi.org/10.1016/j.foreco.2015.06.014
Kohyama T (1993) Size-structured tree populations in gap-dynamic forest: the forest architecture hypothesis for the stable coexistence of species. J Ecol 81:131–143. https://doi.org/10.2307/2261230
Kovac M, Gasparini P, Notarangelo M, Rizzo M, Canellas I, Fernàndez-de-Una L, Alberdi I (2020) Towards a set of national forest inventory indicators to be used for assessing the conservation status of the habitats directive forest habitat types. J Nat Conserv 53:125747. https://doi.org/10.1016/j.jnc.2019.125747
Krajicek JE, Brinkman KA, Gingrich SF (1961) Crown competition: a measure of density. For Sci 7:35–41
Lappi J, Bailey RL (1987) Estimation of the diameter increment function or other tree relations using angle-count samples. For Sci 33:725–739. https://doi.org/10.1093/forestscience/33.3.725
Larrieu L, Cabanettes A, Brin A et al (2014) Tree microhabitats at the stand scale in montane beech–fir forests: practical information for taxa conservation in forestry. Eur J Forest Res 133:355–367. https://doi.org/10.1007/s10342-013-0767-1
Larsson TB, Angelstam P, Balent G, Barbati A, Bijlsma RJ, Boncina A, Bradshaw R et al (2001) Biodiversity evaluation tools for European forests. Ecol Bull 50:127–139
LaRue E, Hardiman B, Elliott J, Fei S (2019) Structural diversity as a predictor of ecosystem function. Environ Res Lett 14:114011. https://doi.org/10.1088/1748-9326/ab49bb
Lassauce A, Paillet Y, Jactel H, Bouget C (2011) Deadwood as a surrogate for forest biodiversity: meta-analysis of correlations between deadwood volume and species richness of saproxylic organisms. Ecol Ind 11:1027–1039. https://doi.org/10.1016/j.ecolind.2011.02.004
Leibundgut H (1982) European primeval mountain forests. Haupt Publishing, Bern-Stuttgart
Leitgeb E, Reiter R, Englisch M, Lüscher P, Schad P, Feger KH (2012) Forest soils: pictures of common soil types of Austria, Germany, and Switzerland. Wiley, Weinheim (**In German**)
Liang J, Buongiorno J, Monserud RA, Kruger EL, Zhou M (2007) Effects of diversity of tree species and size on forest basal area growth, recruitment, and mortality. For Ecol Manage 243:116–127. https://doi.org/10.1016/j.foreco.2007.02.028
Lindenmayer DB, Margules CR, Botkin DB (2000) Indicators of biodiversity for ecologically sustainable forest management. Conserv Biol 14:941–950. https://doi.org/10.1046/j.1523-1739.2000.98533.x
MacArthur RH, MacArthur JW (1961) On bird species diversity. Ecology 42:594–598. https://doi.org/10.2307/1932254
Mace GM, Norris K, Fitter AH (2012) Biodiversity and ecosystem services: A multilayered relationship. Trends Ecol Evol 27:19–26. https://doi.org/10.1016/j.tree.2011.08.006

Magura TB, Tothmeresz M, Bordan Z (2000) Effects of nature management practices on carabid assemblages (Coleoptera: Carabidae) in a non-native plantation. Biol Conserv 93:95–102. https://doi.org/10.1016/S0006-3207(99)00073-7

Margules C, Pressey RL, Williams PH (2002) Representing biodiversity: Data and procedures for identifying priority areas for conservation. J Biosci 27:309–326. https://doi.org/10.1007/BF02704962

Mayer H (1984) Forests of Europe. Gustav Fischer Publishing, Stuttgart (**In German**)

McElhinny C, Gibbons P, Brack C, Bauhus J (2005) Forest and woodland stand structural complexity: its definition and measurement. For Ecol Manage 218:1–24. https://doi.org/10.1016/j.foreco.2005.08.034

McElhinny C, Gibbons P, Brack C (2006) An objective and quantitative methodology for constructing an Index of stand structural complexity. For Ecol Manage 235:54–71. https://doi.org/10.1016/j.foreco.2006.07.024

McNally R, Parkinson A, Horrocks G, Conole L, Tzaros C (2001) Relationships between terrestrial vertebrate diversity, abundance, and availability of coarse woody debris on south-eastern Australian floodplains. Biol Cons 99:191–205. https://doi.org/10.1016/S0006-3207(00)00180-4

Meyer P, Tabaku V, Lüpke BV (2003) Structure of Albanian primeval Beech forests. Forstwiss Centr 122:47–58. https://doi.org/10.1046/j.1439-0337.2003.02041.x. (**In German**)

Motz K, Sterba H, Pommerening A (2010) Sampling measures of tree diversity. For Ecol Manage 260:1985–1996. https://doi.org/10.1016/j.foreco.2010.08.046

Mucina L, Grabherr G, Wallnöfer S (1993) Austrian plant societies III, forests and bushes. Fischer Publishing, Jena (**In German**)

Müller J, Bussler H, Utschick H (2007) How much dead wood does a forest need? A scientifically based concept against species loss in dead wood cenoses. Nat Landsch 39:165–170 (**In German**)

Müller J, Pöllath J, Moshammer R, Schröder B (2009) Predicting the occurrence of Middle Spotted Woodpecker *Dendrocopos medius* on a regional scale, using forest inventory data. For Ecol Manage 257:502–509. https://doi.org/10.1016/j.foreco.2008.09.023

Nascimbene J, Marini L, Motta R, Nimis PL (2008) Influence of tree age, tree size and crown structure on lichen communities in mature Alpine spruce forests. Biodivers Conserv 18:1509. https://doi.org/10.1007/s10531-008-9537-7

Neumann M, Starlinger F (2001) The significance of different indices for stand structure and diversity in forests. For Ecol Manage 145:91–106. https://doi.org/10.1016/S0378-1127(00)00577-6

Newbold T, Hudson LN, Hill SLL, Contu S, Lysenko I, Senior RA, Börger DJ et al (2015) Global effects of land use on local terrestrial biodiversity. Nature 520:45–50. https://doi.org/10.1038/nature14324

Noss RF (1990) Indicators for monitoring biodiversity: a hierarchical approach. Conserv Biol 4:355–364. https://doi.org/10.1111/j.1523-1739.1990.tb00309.x

NW-FVA (2008) Results of applied science on European Beech. Universitätsverlag Göttingen, Göttingen (**In German**)

Olsgard F, Brattegard T, Holthe T (2003) Polychaetes as surrogates for marine biodiversity: lower taxonomic resolution and indicator groups. Biodivers Conserv 12(5):1033–1049

Parkes D, Greame N, Chea D (2003) Assessing the quality of native vegetation: the 'habitat hectares' approach. Ecol Manag Restor 4:29–38. https://doi.org/10.1046/j.1442-8903.4.s.4.x

Pastor J, Moen RA, Cohen Y (1997) Spatial heterogeneities, carrying capacity, and feedbacks in animal-landscape interactions. J Mammal 78:1040–1052. https://doi.org/10.2307/1383047

Pereira HM, Navarro LM, Santos Martins I (2012) Global biodiversity change: the bad, the good, and the unknown. Annu Rev Environ Resour 37:25–50. https://doi.org/10.1146/annurev-environ-042911-093511

Plantureux S, Villerd J, Amiaud B, Taugourdeau S, Bockstaller C (2011) Selection of indicators by machine learning: application to estimate permanent grassland plant richness. General Meeting of the European Grassland Federation. https://hal.inrae.fr/hal-02749822. Accessed 20 Jan 2021.

Podrázský V, Martiník A, Matnjka K, Viewegh J (2014) Effects of Douglas-fir (*Pseudotsuga menziesii*) on understory layer species diversity in managed forests. J for Sci 60:263–271

Posch B, Oitzinger G, Gruber G (2008) Field work manual for BR Vienna Woods core area monitoring. Biosphere Reserve Vienna Woods Management, Vienna (**In German**)

Pretsch H (2002) Basics of forest growth research. Springer Spektrum, Berlin (**In German**)

Pommerening A (2002) Approaches to quantify forest structure. Forestry 75:305–324. https://doi.org/10.1093/forestry/75.3.305

Ponge JF (2003) Humus forms in terrestrial ecosystems: a framework to biodiversity. Soil Biol Biochem 35:935–945. https://doi.org/10.1016/S0038-0717(03)00149-4

Purvis A, Hector A (2000) Getting the measure of biodiversity. Nature 405:212–219. https://doi.org/10.1038/35012221

Quinlan JR (1986) Induction of decision trees. Mach Learn 1:81–106. https://doi.org/10.1007/BF00116251

R Core Team (2020): R: a language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/. Accessed 20 Jan 2021.

Reineke LH (1933) Perfecting a standing density Index for even-aged forests. J Agric Res 46:627–638

Reimoser F (2003) Steering the impacts of ungulates on temperate forest. J Nat Conserv 10:243–252. https://doi.org/10.1078/1617-1381-00024

Rondeux J, Sanchez C (2010) Review of indicators and field methods for monitoring biodiversity within national forest inventories: core variable dead wood. Environ Monit Assess 164:617–630. https://doi.org/10.1007/s10661-009-0917-6

Sala OE, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, Huber-Sanwald E et al (2000) Global biodiversity scenarios for the year 2100. Science 287:1770–1774. https://doi.org/10.1126/science.287.5459.1770

Salmon S, Mantel J, Frizzera L, Zanella A (2006) Changes in humus forms and soil animal communities in two developmental phases of Norway spruce on an acidic substrate. For Ecol Manage 237:47–56. https://doi.org/10.1016/j.foreco.2006.09.089

Salmon S, Artuso N, Frizzera L, Zampedri R (2008) Relationships between soil fauna communities and humus forms: Response to forest dynamics and solar radiation. Soil Biol Biochem 40:1707–1715. https://doi.org/10.1016/j.soilbio.2008.02.007

Schachinger A (1934) The Vienna woods. Verein für Landeskunde und Heimatschutz von Niederösterreich und Wien, Vienna (**In German**)

Schäfer M, Schauermann J (1990) The soil fauna of beech forests: comparison between a mull and a moder soil. Pedobiologia 34:299–314

Schmidt M, Ellenberg H, Heuveldop J, Kriebitzsch WU, Oheimb GV (2002) Major factors impacting herbal species diversity of forests. Treffpunkt Biol Vielfalt II:113–118 (**In German**)

Sdino R (1996) The Crown Competition factor compared with other Stand Density indices. Master thesis, University of Life Sciences Vienna **(In German)**

Shannon C, Weaver W (1949) The mathematical theory of communication. University of Illionois Press, Urbana and Chicago. p. 125

Sterba H (1987) Estimating potential density from thinning experiments and inventory data. For Sci 33:1022–1034. https://doi.org/10.1093/forestscience/33.4.1022

Sterba H (2008) Diversity indices based on angle count sampling and their interrelationships when used in forest inventories. Forestry 8:587–597. https://doi.org/10.1093/forestry/cpn010

Storch F, Dormann CF, Bauhus J (2018) Quantifying forest structural diversity based on large-scale inventory data: a new approach to support biodiversity monitoring. J. Ecosyst. Ecogr. 5:34. https://doi.org/10.1186/s40663-018-0151-1

Svensson JS, Jeglum JK (2001) Structure and dynamics of an undisturbed old-growth Norway spruce forest on the rising Bothnian coastline. For Ecol Manage 15:67–79. https://doi.org/10.1016/S0378-1127(00)00697-6

Tabaku V (1999) Structure of primeval beech forests in Albania in comparison with German beech natural forest reserves and commercial forests. Cuvillier Publishing, Göttingen (**In German**)

Tittensor DP, Walpole M, Hilll SLL, Boyce DG, Britten GL, Burgess ND, Butchart SH et al (2014) A midterm analysis of progress toward international biodiversity targets. Science 346:241–244. https://doi.org/10.1126/science.1257484

Turnbull LA, Isbell F, Purves DW, Loreau M, Hector A (2016) Understanding the value of plant diversity for ecosystem function through niche theory. Philos Trans R Soc B 283:20160536. https://doi.org/10.1098/rspb.2016.0536

Uliczka H, Angelstam P (1999) Occurrence of epiphytic macrolichens in relation to tree species and age in managed boreal forest. Ecography 22:396–405. https://doi.org/10.1111/j.1600-0587.1999.tb00576.x

Ulyshen MD (2011) Arthropod vertical stratification in temperate deciduous forests: Implications for conservation-oriented management. For Ecol Manage 261:1479–1489. https://doi.org/10.1016/j.foreco.2011.01.033

Venables WN, Ripley BD (2002) Modern applied statistics with S. Springer, New York

Vospernik S, Sterba H (2016) Do competition-density rule and self-thinning rule agree? Ann for Sci 72:379–390. https://doi.org/10.1007/s13595-014-0433-x

Waldron A, Miller DC, Redding D, Mooers A, Kuhn TS, Nibellink N, Roberts JT et al (2017) Reductions in global biodiversity loss predicted from conservation spending. Nature 551:364–367. https://doi.org/10.1038/nature24295

Williams LJ, Paquette A, Cavender-Bares J, Messier C, Reich PB (2017) Spatial complementarity in tree crowns explains overyielding in species mixtures. Nat Ecol Evol 1:0063. https://doi.org/10.1038/s41559-016-0063

Winter S, Möller G (2008) Microhabitats in Lowland Beech Forests as monitoring tool for nature conservation. For Ecol Manage 255:1251–1261. https://doi.org/10.1016/j.foreco.2007.10.029

Winter S, Chirici G, McRoberts RE, Hauk E, Tomppo E (2008) Possibilities for harmonizing national forest inventory data for use in forest biodiversity assessments. Int J Environ Res Public Health 81:33–44. https://doi.org/10.1093/forestry/cpm042

Yachi S, Loreau J (2007) Does complementary resource use enhance ecosystem function? A model of light competition in plant communities. Ecol Lett 10:54–62. https://doi.org/10.1111/J.1461-0248.2006.00994.X

Yoccoz NG, Nichols JD, Boulinier T (2001) Monitoring of biological diversity in space and time. Trends Ecol Evol 16:446–453. https://doi.org/10.1016/S0169-5347(01)02205-4

Zhu R, Zeng D, Kosork MR (2015) Reinforcement learning trees. J Am Stat Assoc 110:1770–1784. https://doi.org/10.1080/01621459.2015.1036994

Zheng LT, Chen HY, Yan ER (2015) Tree species diversity promotes litterfall productivity through crown complementarity in subtropical forests. J Ecol 4:1852–1861. https://doi.org/10.1111/1365-2745.13142