Remote Sensing of Environment 186 (2016) 678–686

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Statistical inference for forest structural diversity indices using airborne laser scanning data and the k-Nearest Neighbors technique

Matteo Mura [a], Ronald E. McRoberts [b,*], Gherardo Chirici [c], Marco Marchetti [a]

[a] *Department of Bioscience and Territory, University of Molise, Pesche, Italy*
[b] *Northern Research Station, U.S. Forest Service, Saint Paul, MN, USA*
[c] *Department of Agricultural, Food and Forestry Systems, University of Florence, Florence, Italy*

ARTICLE INFO

Article history:
Received 3 June 2016
Received in revised form 6 September 2016
Accepted 14 September 2016
Available online 11 October 2016

Keywords:
Bootstrap variance
Optimization
Multivariate

ABSTRACT

Forest structural diversity plays a major role for forest management, conservation and restoration and is recognized as a fundamental aspect of forest biodiversity. The assessment, maintenance and restoration of a diversified forest structure have become major foci in the effort to preserve forest ecosystems from loss of biological diversity. However, the assessment of forest biodiversity is difficult because it involves multiple components and is characterized using multiple variables. The objective of the study was to develop a methodological approach for predicting, mapping, and constructing a statistical inference for a multiple-variable index of forest structural diversity. The method included three key components: (i) use of the k-Nearest Neighbors (k-NN) technique, field plot data, and airborne laser scanning metrics to predict multiple forest structural diversity variables simultaneously, (ii) incorporation of multiple diversity variable predictions into a single index, and (iii) construction of a statistically rigorous inference for the population mean of the index. Three structural diversity variables were selected to illustrate the method: growing stock volume and the standard deviations of tree diameter at breast-height and tree height. Optimization of the k-NN technique produced mean relative deviations less in absolute value than 0.04 for predictions for each of the three structural diversity variables, R^2 values between 0.50 and 0.66 which were in the range of values reported in the literature, and a confidence interval for the population mean of the index whose half-width was approximately 5% of the mean. Finally, the spatial pattern depicted in the resulting map of forest structural diversity for the study area contributed to validating the proposed method.

1. Introduction

1.1. Forest structural diversity

Forest structural diversity has been recognized as a fundamental component of forest biodiversity assessment and monitoring (Chirici et al., 2011). Biodiversity, from both general ecological and applied forestry perspectives, is characterized by a larger number of plant and animal communities sharing a common multidimensional space of habitats and niches and making greater use of available resources (McElhinny et al., 2005). Loss of these habitats and niches triggers a loss of biodiversity (Heino et al., 2009; Michel and Winter, 2009). Forests and wooded lands are the richest ecosystems from biological and genetic perspectives (Holdridge, 1947, 1967; Dinerstein et al., 1995), with anthropogenic activities constituting the main causes of the loss of forest biodiversity worldwide (Foley et al., 2005). Thus, the assessment, maintenance and the restoration of forest structural diversity have become major foci in the effort to preserve forest ecosystems from loss of biological diversity.

Forests are complex and adaptive systems (Puettmann et al., 2009) and given the complexity of the biotic and abiotic interactions, compositional, structural, and functional attributes are all involved in the assessment of forest diversity. However, functional attributes describing cycles of mass and energy among the components can be only assumed and only rarely verified, while compositional aspects are rarely measured in the field; thus lack of data and failure to understand fully mechanisms underlying ecophysiological processes cause assessments to become approximate, possibly subjective, and applicable only for isolated cases (Chirici et al., 2011).

In addition, often the only data available to investigate forest diversity at large scales are from national forest inventories (NFI) (Chirici et al., 2011; Pommerening, 2002) for which the primary objective has traditionally been to quantify the amount and extent of forest woody resources while seeking a compromise between the precision of estimates and limited financial resources. This compromise leads to limiting acquisition of field data to information for trees that satisfy size thresholds in the form of minimum diameters at breast-height (DBH), to information for species that contribute most in terms of woody volume, but

---

* Corresponding author.
E-mail address: rmcroberts@fs.fed.us (R.E. McRoberts).

http://dx.doi.org/10.1016/j.rse.2016.09.010
0034-4257/Published by Elsevier Inc.

omission of important ecological information for herbs, shrubs, animals, habitat trees and smaller and younger trees that are important for the functional dynamics of an ecosystem.

However, the structural diversity attributes of the macro component of tree communities gathered by NFIs are objective, reliable, and easy to calculate and understand when compared to more complex indices relying on functional and compositional aspects which offer little useful information for non-expert policy-makers (Branquart and Latham, 2007). In their review which cites studies ranging geographically from boreal to tropical forests, McElhinny et al. (2005) assert that forest structure is more relevant than composition for biodiversity assessments; the rationale is that more diverse stand structures have more niches and, therefore, support more species which results in greater diversity and more efficient use of available resources.

Unmanaged forests tend to have greater structural heterogeneity than managed forests, can better resist the effects of influential internal and external adverse factors (Puettmann et al., 2009), and have greater resilience than managed forests. Moreover, measures of forest structural diversity are characterized by attributes that are judged indispensable for assessing forest diversity. They are reliable in producing objective, consistent and precise results (Uotila et al., 2002; Smith and Theberge, 1987; Liira et al., 2007); they are widely available and easy to calculate (Liira et al., 2007; Bartha et al., 2006); over time they respond to changes in forest dynamics (Angermeier and Karr, 1994); and they are appropriate for assessment at multiple scales (Bartha et al., 2006). Therefore, structural attributes may be reasonably assumed to constitute a reliable basis for objective assessments of forest diversity. Further, from a management perspective, maps of the spatially explicit patterns of structural diversity are of great use for locating hot spots where greater biodiversity is likely to occur, for assisting managers in planning adequate preservation strategies, and for assisting conservationists in prioritizing areas for biodiversity-oriented studies.

1.2. Forest structural diversity measures

Multiple measures of forest structural diversity have been proposed and evaluated. Lexerød and Eid (2006) evaluated eight diameter diversity measures for forest management purposes; Pommerening (2002) evaluated eight measures for habitat functions and forest management planning; Latifi (2011) considered multiple categories of measures that can be estimated using remotely sensed data; and Neumann and Starlinger (2001) evaluated 11 measures for assessing the effects of air pollution. Staudhammer and LeMay (2001) and Müller and Vierling (2014) both noted that of the many measures, the most commonly used include diameter, height (H), or both. With respect to particular measures, Staudhammer and LeMay (2001) reported that the Shannon index performed well for ranking spatial areas with respect to degree of diversity, but Lexerød and Eid (2006) reported that the Gini index was superior for boreal forest planning purposes. The relevant conclusions from the literature are that a multitude of measures of forest structural diversity are feasible, but that prospects for a globally superior measure are unlikely. Further, whereas the vast majority of proposed measures address only a single component of diversity such as height, diameter, or spatial location, diversity encompasses multiple components.

1.3. Objectives

The objective of the study was to develop a methodological approach for predicting, mapping, and constructing a statistical inference for a multiple-variable index of forest structural diversity. The proposed approach relies on prediction of forest structural diversity variables using airborne laser scanning (ALS) metrics (Lim et al., 2003; Zimble et al., 2003; Wulder et al., 2008; Mura et al., 2015) and features three innovative components: (i) use of the multivariate, non-parametric k-Nearest Neighbors technique (k-NN) to predict multiple forest structural diversity response variables simultaneously, (ii) development of a multiple-variable index that integrates any combination of particular single-variable forest structural diversity variables, and (iii) statistically rigorous inference for the population mean of the multiple-variable index. The k-NN technique is well-suited for this approach because it permits simultaneous prediction of multiple response variables, is not constrained by distributional assumptions, and is well-documented for use with forest inventory data (Tomppo et al., 2008). A bootstrap resampling technique was used to estimate the uncertainty of the estimated mean of the multiple-variable structural diversity index. Although the primary study objective was methodological, the approach is illustrated for a study area in Molise, Italy. Of importance, the selected forest structural diversity variables are intended to be illustrative only rather than definitive because relevant forest structural diversity variables vary for each application and study area. Nevertheless, the general approach consisting of the multivariate k-NN prediction technique, the multiple-variable index, and the inferential approach is applicable for any application and study area.

2. Materials and methods

2.1. Study area

The study area included 36,360 ha in the southwestern part of Molise Region in central Italy (Fig. 1). Approximately 56% of the area, corresponding to 20,518 ha, is covered by forests of which 60% is deciduous oaks (*Quercus cerris* L., *Quercus pubescens* Willd), 18% is hop hornbeam (*Ostrya carpinifolia* Scop.), 9% is beech (*Fagus sylvatica* L.), 7% is evergreen holm oak (*Quercus ilex* L.), 4% is hygrophilus forest, and the remainder is species of <1% each. The oak and hop hornbeam forests are mainly privately-owned and managed using a coppice with standards system characterized by rotation ages between 18 and 25 years, cuts of size 1–2 ha, and 100–200 residual standards/ha. Conversely, the beech forests are unmanaged and now have structures approaching natural, old-growth forest status.

2.2. Field data

The study area was tessellated into 437 hexagons, each with area of 1 km^2, and two-phase tessellation stratified sampling (TSS) was conducted (Chirici et al., 2015). In the first phase, a point was randomly selected in each hexagon and classified as “forest” or “non-forest” based on the Italian NFI definition of at least 10% tree cover, minimum area of 0.5 ha, and potential height of at least 5 m at maturity (Gasparini et al., 2013). The attribution of a point as “forest” or “non-forest” was based on interpretation of high-resolution aerial ortho-photography; of the 437 points, 197 were classified as “forest” (Fig. 1). In the second phase, a sampling rate of approximately 30% was applied to the 197 points classified as a “forest” to randomly select 62 points to be visited in the field during years 2009–2011 (Fig. 1). The plot configuration consisted of a circular plot of 13-m radius with measurement of all trees that satisfied the Italian NFI minimum DBH threshold of 9.5 cm (Gasparini et al., 2013). Height was measured for a sub-sample of approximately 10 trees per plot that included the three largest trees, the five trees nearest the plot center, and two trees selected from less frequently observed species and diameter classes. For the remaining trees, H was predicted using a model of the H-DBH relationship constructed using data for the measured trees.

2.3. Structural diversity index (SDI)

An index of forest structural diversity is formulated that incorporates multiple features suggested in the literature. First, the index incorporates multiple forest structural diversity variables as suggested by Neumann and Starlinger (2001), Loidi (1994, pp. 17–30), and Merganič et al. (2012). Second, the index includes variances of vertical and horizontal structure as suggested by Jaehne and Dohrenbusch (1997). Third, as suggested by Staudhammer and LeMay (2001), the