Introduction and evaluation of possible indices of stand structural diversity

Christina Lynn Staudhammer and Valerie Marie LeMay

Abstract: Stand structural diversity is useful in forecasting growth and can be indicative of overall biodiversity. Many variables that indicate structural diversity can be measured. However, species, diameter, and height are commonly measured and indicate changes in vertical and horizontal stand structure. Indices based on the distribution of basal area per hectare by diameter, height, and species were derived and evaluated by applying them to simulated and actual data sets with a wide variety of stand structures. Extending the Shannon index of diversity to proportions by species, diameter, and height resulted in reasonable results with more diverse structures having higher values. However, diameter and height ranges must be divided into classes to use these indices. A new index based on the variances of the target stand, relative to the variance of a uniformly distribution stand, showed similar diversity measures to that of the Shannon index, without the need for dividing the diameter and height data into classes. Examination of these indices for use in growth and yield modelling of complex stands is needed.


Introduction

Measures of stand structural diversity are important for predicting future stand growth. Oliver and Larson (1996) indicated that a variety of patterns of growth are related to structural complexity. High diversity is associated with stands where there are multiple tree species and sizes (Buongiorno et al. 1994). Stand structure is also an important element of stand biodiversity (MacArthur and MacArthur 1961; Willson 1974; Franzreb 1978; Temple et al. 1979; Aber 1979; Ambuel and Temple 1983; Freemark and Merriam 1986). For forested ecosystems, structural diversity can indicate overall species diversity (Kimmins 1997), as shown in research on avian and insect diversity (Whittaker 1972; Franzreb 1978; Aber 1979; Temple et al. 1979; Recher et al. 1996; Moen and Gutierrez 1997). Managing forests for biodiversity may be accomplished by managing for structural diversity (Önal 1997).

Diversity indices have gained wide acceptance in forestry as quantitative measures of species diversity (Swindel et al. 1984; McMinn 1992; Silbaugh and Betters 1995; for a thorough review of diversity indices, see Magurran 1988). Perhaps the most commonly used index is Shannon’s index (Shannon and Weaver 1949; also called the Shannon–Weiner, or the Shannon–Weaver index), which is based on the probability that an individual picked at random from an infinitely large community will be a certain species. The more uncertainty one has about the species of an individual, the higher the diversity of the community. Shannon’s index, $H'$, is defined as follows:

$$H' = -\sum_{i=1}^{S} p_i \ln p_i$$

where $p_i$ is the proportion of individuals in the $i$th species, and $S$ is the number of species. The proportion of a species has been based on a variety of variables including: number of individuals (Franzreb 1978; Swindel et al. 1991; Niese and Strong 1992; Condit et al. 1996), basal area (McMinn
For all of these applications of indices, continuous variables were grouped into classes to calculate proportions. Information about the distributions was lost, and class limits were somewhat arbitrary. Lähde et al. (1999) used three diameter groups (2–10, 11–25, and >25 cm). MacArthur and MacArthur (1961) altered the separation into horizontal layers until three layers were selected. Wikström and Eriksson (2000) used 5-cm diameter classes. Gove et al. (1995) used 5-cm (2 in.) classes, whereas Solomon and Gove (1999) used 2.5-cm (1 in.) classes for diameter rarity profiles. Also, some of the variables used, such as the volume of coarse woody debris or the crown ratio, are difficult to measure and may not always available from forest surveys.

The objective of this study was to develop and evaluate possible indices of stand structural diversity that might be used to classify stands into structural classes, examine structural changes over time, and improve growth predictions. Maximal values for the stand structural diversity indices were based on (i) an even distribution of basal area per hectare (uniform distribution) over a wide size range and (ii) a large number of species, to indicate a wide variety of species and sizes, defined as stand diversity by Lähde et al. (1999). An even distribution was considered a desirable characteristic for species diversity by Pielou (1975) and was applied here for size diversity. Basal area per hectare was used instead of stems per hectare to better represent resource use, with larger trees having more influence (as suggested by LeMay et al. (1997) and Solomon and Gove (1999)). Minimum diversity was similarly defined as one species and size, distributed as a spike at a single point. Diameter outside bark at breast height (DBH; 1.3 m above ground) and total tree height were used to indicate variety in tree size. Other variables could be used; however, Lähde et al. (1999) argued that tree size distribution can be used to indicate stand structure, as it is unambiguous and readily measured. Several indices to indicate differences in species, height, and diameter distributions were derived and evaluated using simulated stand data and also applied to actual stand data from the University of British Columbia’s Malcolm Knapp Research Forest (MKRF) located on coast of southern British Columbia (BC). Of the indices derived, the two most promising indices are presented in this paper.

Description of structural indices

Desirable characteristics

Desirable characteristics that were considered in developing structural indices were that the index should have the following characteristics.

(1) They should equally emphasize horizontal and vertical diversity, since both measures contribute to diversity by creating a variety of habitats. Unequal weighting may be more appropriate for some applications of a stand structural diversity index. For example, bird species may respond more to horizontal rather than vertical variation. However, equal weight of horizontal and vertical variation may be justified as both size measures contribute to changes in productivity and varieties of habitat.
They should not rely on segregating the continuous DBH and height measures into discrete classes. Although class limits could be arbitrarily set for continuous variables, an index that fully utilizes the size information is preferable.

They should be tree size insensitive. A population with small-DBH trees evenly distributed over a 10-cm range should have the same index value as one with large-DBH trees over a 10-cm range. An index with this property could be used to compare stands of different ages.

Given an even distribution, they should be higher if the range of DBH/height is larger, and (or) the number of species is greater. This is similar to Pielou’s (1975) second criterion for species diversity in that, given two completely even communities, the one with more species should be assigned greater diversity. For the continuous variables DBH and height, this translates into higher ranges.

They should be the same for a bimodal population distributed evenly in two separate and distinct canopy layers, as for a unimodal even distribution with a range equal to the sum of the two ranges for the bi-modal distribution.

These characteristics were considered in deriving the indices presented.

Shannon’s index extended to diameter, height, and species

Shannon’s index was applied for species and size diversity by grouping the DBH and height values into classes, as used by other authors. Two different indices were then derived:

1. Post-hoc method: The proportion of basal area per hectare by DBH classes was used in eq. 1 ($H'_D$). The calculation was repeated for height and for species ($H'_H$ and $H'_S$, respectively). The three indices were then averaged to maintain a scale similar to the original index ($H'_{D+S+H}$).

2. Combination method: The proportion of basal area per hectare in each DBH–height–species combination was used to calculate a single index using eq. 1 ($H_{DHS}$).

These two indices measure richness (the number of classes) and evenness of stand structural diversity, giving equal weight to horizontal, vertical, and species diversity. The Shannon index also has been shown to attain the highest values when distributions are perfectly even (i.e., uniform), which can be extended to these indices. For the post-hoc method, the maximum value possible is equal to average of the logarithms of the numbers of species, DBH classes, and height classes. For the combination method, the maximum value is equal to the logarithm of the number of species–DBH–height classes. Both indices are insensitive to tree size, as indices are weighted only on the proportion of basal area occurring in a particular class. However, size variables must be placed into classes to calculate proportions, with some loss of information. A change in class boundaries or an increase in the number of classes would invariably change the value of the indices. As noted by Pielou (1975, p. 8), the combination method applied to hierarchial classes does result in the property that the diversity based on the three-way classification can be subdivided meaningfully into the three separate classes. Pielou uses the example of classifying by genus and species, where the two-way classification can be subdivided into diversity of genus, and diversity of species within genus, that sum to the overall diversity of genus or species. However, no logical hierarchy is represented by the use of species, DBH, and height in the three-way classification.

Structure index based on variance (STVI)

Another index of structural diversity was derived by comparing the variance of the basal area distribution for the target stand to the variance of the theoretically maximally diverse stand. The empirical variance for DBH or height was calculated by

$$ S^2 = \sum_{i=1}^{n} w_i \times (x_i - \bar{x})^2 $$

where $x_i$ is DBH, or height, $\bar{x}$ is the mean of DBH or height, $w_i$ is the basal area per hectare represented by the $i$th tree in the sample plot, and $n$ is the number of trees in a sample plot. The variance of a univariate uniform distribution is given by

$$ S_u^2 = \frac{(b-a)^2}{12} $$

The maximum possible variance of a distribution occurs when the distribution is maximally bimodal, when half the basal area is at $a$ and half the basal area is at $b$. For this basal area distribution, the variance is

$$ S_{max}^2 = \left[ \frac{1}{2} \left( \frac{a+b}{2} - a \right)^2 + \frac{1}{2} \left( \frac{b-a}{2} \right)^2 \right] = \frac{(b-a)^2}{4} $$

The variances for the target stand, the maximally diverse population (uniform distribution) and the bimodal distribution (maximum variance) were used to develop an index. For convenience, the index should have a value of 1.0 for the most diverse stand, when the variance for the empirical distribution is equal to that of the uniform distribution. Also, the index should be near zero when the empirical distribution has zero variance (all values are one DBH or height), or when the variance is very close to the maximally bimodal distribution. For variances between the two ranges, from zero variance to the variance of a uniform distribution, and from the variance of a uniform distribution to that of a bimodal distribution, the value of the index should vary between 0 and 1. The equation developed to calculate the diversity index, as shown for DBH ($STVI_{DBH}$ for a species $k$), was

$$ STVI_{DBH} = \left\{ \begin{array}{ll}
\frac{S^2_{DBH_{max}} - S^2_{DBH}}{S^2_{DBH_{max}}} & \text{when } S^2_{DBH} \leq S^2_{DBH_{max}} \\
1 - \frac{m \times S^2_{DBH} - S^2_{DBH_{max}}}{S^2_{DBH_{max}} - S^2_{DBH_{max}}} & \text{when } S^2_{DBH} > S^2_{DBH_{max}}
\end{array} \right. $$

$S^2_{DBH}$ is the basal area per hectare represented by the $k$th tree in the sample plot, and $m$ is the number of trees in the sample plot.
where $S_{DBH}^2$ is the variance of DBH for species $k$ in the target stand, $S_{DBH}^2_{min}$ is the variance when DBH values occur uniformly over the range from $a$ to $b$, $S_{DBH}^2_{max}$ is the variance when half the DBH values occur at each of the extreme ends of the range, $p_1$ and $p_2$ are constants, and $m$ is a constant $\geq 1.0$. The constants $p_1$ and $p_2$ define the shape of the curve relating the value of the index to the sample variance: when $p_1$ (or $p_2$) < 1, the curve is concave upward; when $p_1$ (or $p_2$) > 1, the curve is concave downward (Fig. 1). If $p_1 = p_2 = 1$, then a smooth, continuous function results. The coefficient $m$ controls the value of the index when the distribution is maximally bimodal. When $m = 1$, then the index will be zero for a maximally bimodal distribution; as $m$ gets larger, the index value increases for the maximally bimodal case (Fig. 2).

The values for $p_1$, $p_2$, and $m$ were chosen by placing three constraints on the index to yield certain index values under defined conditions. The index was constrained to equal 0.5 when (i) the variance of the target stand is equal to that of a uniform distribution over half the maximum possible range ($S_{DBH}^2_{min}$) and (ii) the variance of the target stand is equal to that of a bimodal distribution, with half of the values uniformly distributed over the lower quartile, and the other half uniformly distributed over the upper quartile of the maximum possible range ($S_{DBH}^2_{max}$). The index was also constrained to equal 0.1 for the maximum variance for a uniform distribution (STVI$_{DBH}$). Again illustrating this using DBH, the three constraints were

$$
0.5 = \text{STVI}_{DBH} = 1 - \left( \frac{S_{DBH}^2 - S_{DBH}^2_{min}}{S_{DBH}^2_{max}} \right)^{p_1}
$$

$$
0.5 = \text{STVI}_{DBH} = 1 - \left( \frac{S_{DBH}^2 - S_{DBH}^2_{min}}{m \times S_{DBH}^2_{max} - S_{DBH}^2_{min}} \right)^{p_1}
$$

$$
0.1 = \text{STVI}_{DBH} = 1 - \left( \frac{S_{DBH}^2 - S_{DBH}^2_{min}}{m \times S_{DBH}^2_{max} - S_{DBH}^2_{min}} \right)^{p_1}
$$

Based on these constraints, $p_1 \equiv 2.4094$, $p_2 \equiv 0.5993$, and $m \equiv 1.1281$ (see the Appendix for the derivation). To arrive at a measure of structural diversity for species $k$, STVI$_{DBH}$ and STVI$_{height}$ were averaged, with a maximum value of one (uniform for both DBH and height). These were summed over all species in the plot (STVI$_{dbh}$), for a maximum value equal to the number of species.

For a bivariate version of this index, the variance was described by the generalized variance (Johnson and Wichern 1998), and calculated as

$$
S_{DBH, height}^2 = \text{det} \begin{bmatrix}
S_{DBH}^2 & \text{cov}(DBH, height) \\
\text{cov}(DBH, height) & S_{height}^2
\end{bmatrix}
$$

where $S_{DBH}^2$ and $S_{height}^2$ are given in eq. 3, det indicates that the determinant is obtained, and

$$
\text{cov}(dbh, height) = \sum_{i=1}^{n} w_i(DBH_i - \overline{DBH}) \times (height_i - \overline{height})
$$

where DBH and height are averages for DBH and height, respectively. For a bivariate uniform distribution of DBH and height, all values of DBH and height are possible over the two ranges, resulting in a zero covariance. Given the values for maximum and minimum DBH ($a_1$ and $b_1$) and height ($a_2$ and $b_2$), eq. 10 simplifies to

$$
S_U^2 = \text{det} \begin{bmatrix}
(b_1 - a_1)^2 & 0 \\
0 & (b_2 - a_2)^2
\end{bmatrix} = (b_1 - a_1)^2 \times (b_2 - a_2)^2
$$

Similar to the univariate case, the maximum possible variance occurs when one-quarter of the basal area is at each of four extreme points on the DBH–height plane: ($a_1$, $a_2$), ($a_1$, $b_2$), ($b_1$, $a_2$), and ($b_1$, $b_2$). In this situation, the variance of height and of DBH is calculated using eq. 5. As was the case
with the bivariate uniform, the cov(DBH, height)\textsubscript{max} is zero, and the variance is calculated as

\[ S_{\text{max}}^2 = \text{det} \begin{bmatrix} \frac{b_1 - a_1}{4} & 0 \\ 0 & \frac{b_2 - a_2}{4} \end{bmatrix} = \frac{(b_1 - a_1)^2 \times (b_2 - a_2)^2}{16} \]

The bivariate STVI is then

\[ \text{STVI}_{\text{dbh}} = \begin{cases} \left( 1 - \frac{S^2_U - S^2_k}{S^2_U} \right)^p, & \text{when } S^2_U \leq S^2_{\text{dbh}} \\ \left( 1 - \frac{m \times S^2_{\text{max}} - S^2_U}{m \times S^2_{\text{max}} - S^2_{\text{dbh}}} \right)^p, & \text{when } S^2_{\text{dbh}} > S^2_U \end{cases} \]

where \( S^2_k \) is the generalized variance for species \( k \), \( p_1 \) and \( p_2 \) are constants >0, and \( m \) is a constant ≥1.0. As in the univariate case, the powers, \( p_1 \) and \( p_2 \), define the shape of the curve, and the coefficient \( m \) controls the value of the index when the distribution has the maximum variance. Imposing the same constraints on the index as in the univariate case, the same values of \( p_1 \), \( p_2 \), and \( m \) result. An overall measure of diversity for a sample plot was labelled as \( \text{STVI}_{\text{dbh}} \), the sum of the values of all \( \text{STVI}_{\text{dbh}} \) over a plot, with a maximum value equal to the number of species.

Both forms of the STVI indicate the range and evenness of basal area per hectare over tree size. The indices do not rely on combining data into classes, and account for vertical and horizontal diversity equally. Since the indices are based on the variances of the DBH and height distributions, they are insensitive to tree size. Furthermore, providing that \( p_1 \), \( p_2 \), and \( m \) are well chosen, the index value for a bimodal distribution with two distinct ranges would be the same as that of a unimodal distribution with a range equal to the sum of the two ranges. However, species evenness was not accounted for in this proposed index, since all species are given equal weight in calculating the plot value. Modifying the indices by weighting by species proportions could be considered; the maximum value would then be changed from the number of species to one. With the proposed index, the stand with more species would have the higher index value, given even size distributions within each species.

**Evaluation of indices**

Proposed indices were first evaluated using simulated stands, based on different DBH/height distributions, but only one species. Data from the MKRF were used to set DBH and height ranges for the simulated data, and were also used to subsequently introduce species into the evaluation.

**Description of MKRF data**

The MKRF is located in the Coastal Western Hemlock biogeoclimatic (BEC) zone of southwestern BC. The region encompasses low to middle elevations west of the Coastal Mountains and is the wettest BEC zone in BC, receiving 1000–4000 mm of precipitation annually (Meidinger and Pajar 1991). The climate is cool mesothermal, with a mean annual temperature of 8°C and mild winters. The most common species in the forest cover are Douglas-fir (Pseudotsuga menziesii (Mirb.) Franco), western hemlock (Tsuga heterophylla (Raf.) Sarg.), western redcedar (Thuja plicata Donn), amabilis fir (Abies amabilis (Doug.) Forbes), Sitka spruce (Picea sitchensis (Bong.) Carrèrie), and red alder (Alnus rubra Bong.). The main species reaches over 100 cm DBH and over 50 m height.

In 1995, sample plot data were collected using the Ministry of Forests, Vegetation Inventory Sampling Procedures (Resources Inventory Branch 1994). Eighty-two clusters of five plots were systematically located over the research forest using square spacing. Because plots occurred throughout the whole of the MKRF, the data represent a variety of microsites and growing conditions.

For all trees above 2 cm DBH, the species, DBH (cm) and tree class (live or dead, standing or fallen) were obtained. For a subset of trees, the total height (m), crown class (dominant, codominant, intermediate, or suppressed), and height to live crown (m) were also measured. Height was measured on only 763 trees (44%). Height prediction models were developed to estimate the remaining heights from measured DBH and plot variables (Staudhammer and LeMay 2000).

**Simulated tree data**

Simulated stands were used to obtain well-defined cases of different DBH–height distributions, which may be difficult to find in natural data sets. The cases were based on Oliver and Larson (1996) who noted that (i) the distribution of stems per hectare of single-cohort, single-species stands has been shown to follow a normal distribution, whereas well-differentiated stands can be bimodal or skewed; (ii) a skewed distribution may result when intermediate or suppressed trees die readily, or in very old or very young stands; and (iii) in mixed species stands, where tolerant suppressed trees form lower strata, a bimodal distribution may result. Assuming a single species or a pooling of species, 10 stands were simulated with varying structural diversity.

The maximally diverse structure was first chosen (case 1) with a uniform distribution of basal area per hectare over DBH with the range \([r_{11}, r_{12}]\) and over height with the range \([r_{21}, r_{22}]\). This represents a reverse-J distribution of stems per hectare. Ranges were set from 0.1 to 120 cm for DBH, and from 0.1 to 60.0 m for height, based on those found in the MKRF data. Not all generated DBH/height ratios were biologically possible. Since 90% of the MKRF data had DBH/height ratios in the range of 0.8 to 2.4 cm/m, restrictions were applied using these values. The univariate distributions of basal area per hectare by DBH and height were, therefore, uniformly distributed, but the joint distribution of DBH and height did not follow a bivariate uniform distribution (Fig. 3). Cases 2 through 6 were also generated as uniform distributions of basal area per hectare over DBH and over height, but with different ranges. Cases 2 and 3 had two-thirds of the range of case 1, with case 2 representing the upper ends and case 3 representing the lower ends of the DBH and height ranges. Cases 4, 5, and 6 had one-third of the ranges of case 1, representing the upper, middle, and lower ends of the ranges, respectively. Cases 7 and 8 were generated assuming a normal distribution, with a mean equal to case 1.
The variance was set at 20% of the mean for case 7, the more diverse case, and 10% of the mean for case 8. Cases 9 and 10 were bimodal distributions, with half the observations at the upper end and the other half at the lower end of the range. Case 10 was similar to case 9, but had a more limited range of values at the two extremes.

For the uniform distributions and bimodal distributions (cases 1–6 and cases 9 and 10), the SAS function RANUNI (SAS Institute Inc. 1988) was used to generate random numbers, which were scaled to the desired DBH and height ranges. For cases 7 and 8, the SAS function RANNOR was used to generate normally distributed numbers.

To compute the extended Shannon index using the post-hoc method, tree data were classed in 10-cm DBH classes, from 0 to 10 cm through 110 to 120 cm, and $H_a$ was computed. Similarly, tree data were classed into 5-m height ranges...
classes, from 0–5 m through 55–60 m, and \( H_d \) was then computed. The final index was computed as the average of the diameter and height indices and labelled \( H_{d+h} \). For the combination method, the data were classified simultaneously by DBH and height, and \( H_{d+h} \) was computed. The univariate and bivariate STVI were computed, using the maximum and minimum DBH and height values used to generate the 10 cases.

The values for each of the indices were then examined to determine if the indices reflected the diversity of sizes represented in the simulated data. Each case was simulated 100 times to indicate the variability among simulations, and to avoid simulation anomalies.

**Selected plots from the MKRF**

Clusters were selected from the MKRF that closely matched the DBH and height distributions in simulated cases 1–9 based on the basal area distributions for all species pooled together (Table 1). A distribution similar to case 10 was not found in the MKRF data.

The extended Shannon’s indices and STVI indices were calculated on the MKRF clusters. Also, the species were pooled together (Table 1). A distribution similar to case 10 based on the basal area distributions for all species set, such that the number of trees that fell outside of the range was small (DBH from 2 to 150 cm and height from 1.3 to 60 m). The STVI were calculated over all species for DBH and for height separately, and then averaged. The indices were evaluated based on the ability to reflect the diversity of the DBH, height, and species diversity of the cluster.

**Evaluation results and discussion**

Although only the simulations means are given in Table 2, the values for the extended Shannon’s indices were quite similar over all 100 simulations, with slightly higher variability for the combination method and for the normally distributed cases (cases 7 and 8; all standard deviations less than 0.07). The extended Shannon’s index, combination method, tended to be larger than the post-hoc index. This was expected, as classifying by DBH and height simultaneously resulted in more classes, and, therefore, larger values. The post-hoc index resulted in the highest value for case 1, in which the DBH and height distributions were both uniform over a wide range (Fig. 3). Because of the restrictions placed on obtaining only reasonable values of height and DBH combinations, the number of combined DBH–height classes was largest for case 7, the normal distribution with a wide range (Fig. 3). Therefore, the highest value for the combination method was obtained for case 7, with case 1

---

**Table 1.** Descriptive statistics of selected Malcolm Knapp Research Forest plots.

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Diversity; distribution</th>
<th>No. of species</th>
<th>Mean DBH (cm)</th>
<th>Minimum DBH (cm)</th>
<th>Maximum DBH (cm)</th>
<th>Mean Height (m)</th>
<th>Minimum Height (m)</th>
<th>Maximum Height (m)</th>
<th>SD DBH (cm)</th>
<th>SD Height (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Very diverse; uniform</td>
<td>3</td>
<td>52.80</td>
<td>8.4</td>
<td>101.0</td>
<td>34.67</td>
<td>5.9</td>
<td>51.5</td>
<td>25.32</td>
<td>11.70</td>
</tr>
<tr>
<td>2</td>
<td>Moderate diversity; uniform; large</td>
<td>3</td>
<td>84.80</td>
<td>31.7</td>
<td>180.5</td>
<td>47.15</td>
<td>26.8</td>
<td>66.2</td>
<td>42.99</td>
<td>11.13</td>
</tr>
<tr>
<td>3</td>
<td>Moderate diversity; uniform; small</td>
<td>3</td>
<td>34.55</td>
<td>6.5</td>
<td>61.9</td>
<td>25.90</td>
<td>4.6</td>
<td>42.9</td>
<td>14.04</td>
<td>8.82</td>
</tr>
<tr>
<td>4</td>
<td>Low diversity; uniform; large</td>
<td>3</td>
<td>58.24</td>
<td>26.3</td>
<td>99.3</td>
<td>32.00</td>
<td>21.0</td>
<td>46.9</td>
<td>17.47</td>
<td>7.33</td>
</tr>
<tr>
<td>5</td>
<td>Low diversity; uniform; medium</td>
<td>3</td>
<td>56.17</td>
<td>4.6</td>
<td>93.5</td>
<td>28.58</td>
<td>4.2</td>
<td>40.8</td>
<td>19.66</td>
<td>6.97</td>
</tr>
<tr>
<td>6</td>
<td>Low diversity; uniform; small</td>
<td>6</td>
<td>12.47</td>
<td>2.5</td>
<td>24.7</td>
<td>10.07</td>
<td>3.2</td>
<td>15.9</td>
<td>6.31</td>
<td>3.79</td>
</tr>
<tr>
<td>7</td>
<td>Moderate diversity; normal</td>
<td>3</td>
<td>44.84</td>
<td>17.8</td>
<td>73.6</td>
<td>29.56</td>
<td>13.0</td>
<td>42.2</td>
<td>12.23</td>
<td>6.07</td>
</tr>
<tr>
<td>8</td>
<td>Low diversity; normal</td>
<td>3</td>
<td>33.74</td>
<td>21.3</td>
<td>44.9</td>
<td>23.46</td>
<td>16.5</td>
<td>30.2</td>
<td>7.27</td>
<td>4.47</td>
</tr>
<tr>
<td>9</td>
<td>Moderate diversity; bimodal</td>
<td>3</td>
<td>95.42</td>
<td>3.7</td>
<td>174.4</td>
<td>42.50</td>
<td>1.3</td>
<td>68.3</td>
<td>54.97</td>
<td>16.74</td>
</tr>
</tbody>
</table>

---

**Table 2.** Average indices over 100 runs for each simulated stand (cases).

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Diversity; distribution</th>
<th>Extended Shannon’s</th>
<th>STVI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Post-hoc ( H_{d+h} )</td>
<td>Combination ( H_{d+h} )</td>
</tr>
<tr>
<td>1</td>
<td>Very diverse; uniform</td>
<td>2.485</td>
<td>2.522</td>
</tr>
<tr>
<td>2</td>
<td>Moderate diversity; uniform; large</td>
<td>2.079</td>
<td>2.115</td>
</tr>
<tr>
<td>3</td>
<td>Moderate diversity; uniform; small</td>
<td>2.079</td>
<td>2.119</td>
</tr>
<tr>
<td>4</td>
<td>Low diversity; uniform; large</td>
<td>1.391</td>
<td>1.436</td>
</tr>
<tr>
<td>5</td>
<td>Low diversity; uniform; medium</td>
<td>1.392</td>
<td>1.438</td>
</tr>
<tr>
<td>6</td>
<td>Low diversity; uniform; small</td>
<td>1.392</td>
<td>1.441</td>
</tr>
<tr>
<td>7</td>
<td>Moderate diversity; normal</td>
<td>1.494</td>
<td>2.676</td>
</tr>
<tr>
<td>8</td>
<td>Low diversity; normal</td>
<td>0.950</td>
<td>1.812</td>
</tr>
<tr>
<td>9</td>
<td>Moderate diversity; bimodal</td>
<td>2.082</td>
<td>2.121</td>
</tr>
<tr>
<td>10</td>
<td>Low diversity; bimodal</td>
<td>0.693</td>
<td>0.693</td>
</tr>
</tbody>
</table>
having the next highest value. Both indices were insensitive
to tree size, in that similar values were given for cases 2 and
3; for cases 4, 5, and 6; and for case 2 versus case 9
(unimodal vs. bimodal), which differed only in the sizes rep-
resented. The limited biomodal (case 10) was given lower
values than the wider range biomodal (case 9) for both indi-
ces. The extended Shannon’s indices appeared to perform
well for a single species.

For the STVI indices, the average indices over the 100
simulations were quite different for the univariate (STVI_{dbh})
versus the bivariate (STVI_{dbh, height}) indices (Table 2), with much
larger values for the univariate approach. For the bivariate
uniform distribution, all combinations of DBH and height
are equally possible resulting in high variances (Table 3),
which are not biologically possible. Simulated cases were
limited to feasible values of DBH and height for every spe-
cies, or Shannon’s index for DBH, height, and species
respectively, the STVI values for DBH and height for every spe-
cies were much lower than the bivariate uniform (Table 5), re-
sulting in much lower variances for the bivariate STVI, as was
the case in analysing the simulated data. The univariate indi-
ces gave results similar to Shannon’s index. Since the maxi-
num for the univariate STVI is equal to the number of
species, the interpretation is relatively simple. For example,
in case 1, the sum of STVI values for DBH over the three
species was 2.061, indicating high DBH diversity. The diver-
sity in height was less, resulting in a value of 1.126. Over
the six species, case 6 indicated very low DBH (0.321) and
height (0.295) diversity.

Overall, the extended Shannon indices ranked the simu-
lated stands in a logical manner and are size invariant. The univariate STVI also performed well, and data do not need
to be divided into arbitrary classes. However, the sampling
properties of this new index are not known. For the post-hoc
extended Shannon’s index, equal weight was given to aver-
aging the DBH, height, and species indices. Similarly, for
the univariate STVI, vertical and horizontal variations were
given equal weight in obtaining the index by species. The re-
resulting indices were then summed over all species to obtain
a maximum value equal to the number of species. Different
weightings could be introduced, depending on the use of the
index. For example, Kangas and Pukkala (1996) used a
weighting formula in combining different variables. Alterna-
atively, the STVI values for DBH and height for every spe-
cies, or Shannon’s index for DBH, height, and species

### Table 3. Average sample variances for simulated cases compared with the variances of the uniform and maximally bimodal distributions.

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Diversity; distribution</th>
<th>$S_{DBH}^2$</th>
<th>$S_{height}^2$</th>
<th>$S_{DBH, height}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Very diverse; uniform</td>
<td>1219</td>
<td>306</td>
<td>4815</td>
</tr>
<tr>
<td>2</td>
<td>Moderate diversity; uniform; large</td>
<td>549</td>
<td>117</td>
<td>2141</td>
</tr>
<tr>
<td>3</td>
<td>Moderate diversity; uniform; small</td>
<td>539</td>
<td>133</td>
<td>2075</td>
</tr>
<tr>
<td>4</td>
<td>Low diversity; uniform; large</td>
<td>133</td>
<td>34</td>
<td>520</td>
</tr>
<tr>
<td>5</td>
<td>Low diversity; uniform; medium</td>
<td>137</td>
<td>34</td>
<td>523</td>
</tr>
<tr>
<td>6</td>
<td>Low diversity; uniform; small</td>
<td>136</td>
<td>34</td>
<td>527</td>
</tr>
<tr>
<td>7</td>
<td>Moderate diversity; normal</td>
<td>143</td>
<td>26</td>
<td>2959</td>
</tr>
<tr>
<td>8</td>
<td>Low diversity; normal</td>
<td>36</td>
<td>8</td>
<td>256</td>
</tr>
<tr>
<td>9</td>
<td>Moderate diversity; bimodal</td>
<td>1768</td>
<td>441</td>
<td>7025</td>
</tr>
<tr>
<td>10</td>
<td>Low diversity; bimodal</td>
<td>3655</td>
<td>912</td>
<td>0</td>
</tr>
<tr>
<td>Uniform</td>
<td>1198*</td>
<td>299</td>
<td>358 203</td>
<td></td>
</tr>
<tr>
<td>Maximally bimodal</td>
<td>3594</td>
<td>897</td>
<td>3223 829</td>
<td></td>
</tr>
</tbody>
</table>

*Values are slightly lower for the uniform than for case 1, and for the maximally bimodal than for case 10, because the variance calculated for the simulated cases was calculated by dividing by $n−1$ (the sample variance) instead of dividing by $n$ (the population variance).
separately, could be retained, rather than being combined into a single value. This would allow for differential weightings by users of the values. If this information was used in harvest modelling, the values could be used as constraints, similar to work by Wikstrom and Eriksson (2000). For use in developing growth prediction equations, the use of a single index to indicate stand structure will result in simpler equations. Also, equal weighting of horizontal and vertical structure is intuitively appealing for modelling changes in stand structure. Testing the indices for a variety of applications is needed to further assess their usefulness in growth modelling, classifying stand structure, and examining structural changes over time.

Conclusions

Several indices of structural diversity were proposed and evaluated using simulated stands, and using data from the MKRF. For all proposed indices, the basal area distribution was used to better represent site occupation by trees. The extensions of Shannon’s index to DBH, height, and species performed well in ranking the structural diversity and indicated species evenness. Of the two extensions to Shannon’s index, the post-hoc method resulted in more information, by indicating diversity in each variable, as well as overall diversity. However, both DBH and height must be divided into arbitrary classes to calculate proportions, and the number of classes used alters the maximum value of the index.

The STVIs, based on the variance of the stand relative to the variance of uniform and bimodal distributions, do not require arbitrary classes. However, the bivariate form of this index was based on a maximum variance, which will not be achieved in a natural setting, where not all combinations of DBH and height values are possible. The univariate form of the STVI gave good results that can be relatively easily interpreted.

Of the indices tested, the post-hoc extended Shannon index or the univariate STVI are recommended. However, possible improvements have been identified and could be investigated.

Table 4. Indices for the MKRF data.

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Diversity; distribution</th>
<th>$H^*_DBH$</th>
<th>$H^*_height$</th>
<th>$H^*_DBH+height$</th>
<th>$S_{DBH}$</th>
<th>$S_{DBH+height}$</th>
<th>STVI$_{DBH}$</th>
<th>STVI$_{height}$</th>
<th>STVI$_{DBH+height}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Very diverse; uniform</td>
<td>2.186</td>
<td>2.091</td>
<td>2.152</td>
<td>2.860</td>
<td>2.061</td>
<td>1.126</td>
<td>1.594</td>
<td>0.085</td>
</tr>
<tr>
<td>2</td>
<td>Moderate diversity;</td>
<td>2.173</td>
<td>1.935</td>
<td>2.172</td>
<td>3.103</td>
<td>1.090</td>
<td>1.984</td>
<td>1.537</td>
<td>0.498</td>
</tr>
<tr>
<td>3</td>
<td>Uniform small</td>
<td>1.748</td>
<td>1.915</td>
<td>0.997</td>
<td>1.553</td>
<td>0.733</td>
<td>0.391</td>
<td>0.562</td>
<td>0.009</td>
</tr>
<tr>
<td>4</td>
<td>Low diversity; uniform</td>
<td>1.732</td>
<td>1.736</td>
<td>0.791</td>
<td>1.420</td>
<td>0.652</td>
<td>0.678</td>
<td>0.630</td>
<td>0.031</td>
</tr>
<tr>
<td>5</td>
<td>Low diversity;</td>
<td>1.059</td>
<td>1.087</td>
<td>1.040</td>
<td>2.747</td>
<td>1.087</td>
<td>0.840</td>
<td>0.515</td>
<td>0.042</td>
</tr>
<tr>
<td>6</td>
<td>Uniform small</td>
<td>0.958</td>
<td>1.117</td>
<td>1.659</td>
<td>1.245</td>
<td>0.321</td>
<td>0.295</td>
<td>0.308</td>
<td>0.000</td>
</tr>
<tr>
<td>7</td>
<td>Moderate diversity;</td>
<td>1.616</td>
<td>1.562</td>
<td>1.011</td>
<td>1.396</td>
<td>1.028</td>
<td>0.697</td>
<td>0.863</td>
<td>0.010</td>
</tr>
<tr>
<td>8</td>
<td>Normal</td>
<td>0.995</td>
<td>1.149</td>
<td>0.981</td>
<td>1.042</td>
<td>0.482</td>
<td>0.172</td>
<td>0.327</td>
<td>0.001</td>
</tr>
<tr>
<td>9</td>
<td>Bimodal</td>
<td>2.141</td>
<td>1.851</td>
<td>1.098</td>
<td>2.497</td>
<td>1.298</td>
<td>1.853</td>
<td>1.575</td>
<td>0.215</td>
</tr>
</tbody>
</table>

Note: The maximum value for case 6 is ln(6) = 1.792 and, for all other cases, is ln(3) = 1.099 based on species alone.

Table 5. Sample variances for data pooled over species and averaged by species for chosen MKRF plots, compared with the variance of a theoretical uniform distribution.

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Diversity; distribution</th>
<th>No. of species</th>
<th>DBH, $S^2_{DBH}$</th>
<th>Height, $S^2_{height}$</th>
<th>DBH + height, $S^2_{DBH+height}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Very diverse;</td>
<td>3</td>
<td>641</td>
<td>351</td>
<td>137</td>
</tr>
<tr>
<td>2</td>
<td>Moderate diversity;</td>
<td>3</td>
<td>1848</td>
<td>1280</td>
<td>124</td>
</tr>
<tr>
<td>3</td>
<td>Uniform small</td>
<td>3</td>
<td>197</td>
<td>104</td>
<td>78</td>
</tr>
<tr>
<td>4</td>
<td>Low diversity;</td>
<td>3</td>
<td>305</td>
<td>293</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>Low diversity;</td>
<td>3</td>
<td>386</td>
<td>154</td>
<td>49</td>
</tr>
<tr>
<td>6</td>
<td>Low diversity;</td>
<td>6</td>
<td>40</td>
<td>47</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>Moderate diversity;</td>
<td>3</td>
<td>149</td>
<td>197</td>
<td>37</td>
</tr>
<tr>
<td>8</td>
<td>Uniform small</td>
<td>3</td>
<td>53</td>
<td>45</td>
<td>20</td>
</tr>
<tr>
<td>9</td>
<td>Bimodal</td>
<td>3</td>
<td>3022</td>
<td>991</td>
<td>280</td>
</tr>
</tbody>
</table>

Note: Uniform = $S^2_{DBH}$ = $S^2_{height}$ = $S^2_{DBH+height}$ = 1825; Maximally bimodal = $S^2_{DBH}$ = $S^2_{height}$ = $S^2_{DBH+height}$ = 5476.
for both indices. Application of the indices to a variety of stand structures is needed to determine the usefulness as proxies for the species, DBH, and height distributions.

References


© 2001 NRC Canada
Appendix

Derivation of $p_1$, $p_2$, and $m$

The expected value of the uniform distribution on the interval from $a$ to $b$ is

$$ E[X] = \frac{b + a}{2} \tag{A1} $$

The expected value of the square of the uniform distribution is

$$ E[X^2] = \frac{b^2 + ba + a^2}{3} \tag{A2} $$

with the variance then equal to

$$ S_U^2 = \left( \frac{b^2 + ba + a^2}{3} \right) \left( \frac{b + a}{2} \right)^2 = \left( \frac{b - a}{2} \right)^2 \tag{A3} $$

For a maximally bimodal distribution with half of the observations at $a$ and the other half at $b$, the expected value is the same as for the uniform (eq. A1), but the expected value of the square of this distribution is

$$ E[X^2] = 0.5b^2 + 0.5a^2 \tag{A4} $$

The variance for the maximally bimodal distribution is therefore

$$ S^2_{max} = (0.5b^2 + 0.5a^2) - \left( \frac{b + a}{2} \right)^2 = \left( \frac{b - a}{2} \right)^2 / 4 \tag{A5} $$

which is three times that of a uniform distribution from $a$ to $b$ (eq. A3). For a uniform distribution with one-half of the range of the uniform from $a$ to $b$, the variance is

$$ S^2_{0.5U} = \left( \frac{(b-a)/2}{2} \right)^2 = \left( \frac{b-a}{2} \right)^2 / 48 \tag{A6} $$

which is one-quarter of the variance of a uniform distribution from $a$ to $b$ (eq. A3). For a distribution separated into two discrete uniform ranges from $a$ to $c$ in the lower quartile of the range $a$ to $b$, and from $d$ to $b$ in the upper quartile of the range $a$ to $b$, the range, $r$, is the same for both regions:

$$ r = \frac{b - a}{4} = b - d = c - a $$

and

$$ c = a + r $$
$$ d = b - r $$

The variance of the distribution is

$$ S_B^2 = (7/48)(b-a)^2 \tag{A7} $$

or 1.75 times that of a uniform distribution from $a$ to $b$. This is derived from the expected value for this distribution given in eq. A1, and the expected value of the square of the distribution given as

$$ E[X^2] = \int_a^c x^2 \left( \frac{0.5}{r} \right) dx + \int_d^b x^2 \left( \frac{0.5}{r} \right) dx $$

$$ \tag{A8} $$

Given these specific types of distributions, constraints on the univariate STVI were used to obtain values for $p_1$, $p_2$, and $m$. Constraining the index to a value of 0.5 when the variance of the distribution was equal to that for one-half the range of the uniform, a value for was obtained by

$$ 0.5 = 1 - \left( \frac{S^2_U - S^2_{0.5U}}{S^2_U} \right)^{p_1} \tag{A9} $$

Since the variance for one-half of the range is one-quarter that of a uniform:

$$ 0.5 = 1 - \left( \frac{S^2_B - S^2_{0.5U}}{S^2_B} \right)^{p_1} \tag{A10} $$

Solving this results in $p_1 = 2.0904$. For the second constraint, when $S^2_B = S^2_B$, the index was constrained to equal to 0.5, as follows:

$$ 0.5 = 1 - \left( \frac{S^2_B - S^2_{max}}{m \times S^2_{max} - S^2_U} \right)^{p_1} \tag{A11} $$

Since $S^2_B = 1.75S^2_U$ and $S^2_{max} = 3S^2_U$, this equation becomes

$$ 0.5 = 1 - \left( \frac{1.75S^2_U - S^2_{max}}{m \times 3S^2_U - S^2_U} \right)^{p_1} \tag{A12} $$

The last constraint was used to set the index to 0.1 when $S^2_k = S^2_{max}$

$$ 0.1 = 1 - \left( \frac{S^2_{max} - S^2_U}{m \times S^2_{max} - S^2_U} \right)^{p_1} \tag{A13} $$

Substituting $S^2_{max} = 3S^2_U$, the equation becomes

$$ 0.1 = 1 - \left( \frac{3S^2_U - S^2_U}{m \times 3S^2_U - S^2_U} \right)^{p_1} \tag{A14} $$

Solving eqs. A10 and A11 simultaneously, $p_2 = 0.5993$ and $m = 1.1281$. © 2001 NRC Canada