

Fuzzy Set Classification of Old-Growth Southern Pine

by
Don C. Bragg¹

Abstract

I propose the development of a fuzzy set ordination (FSO) approach to old-growth classification of southern pines. A fuzzy systems approach differs from traditional old-growth classification in that it does not require a “crisp” classification where a stand is either “old-growth” or “not old-growth”, but allows for fractional membership in the set of old-growth stands. FSO produces a score ranging from 0.0 (highly different from old-growth) to 1.0 (completely residing in the set of old-growth stands). This value can also be interpreted as “apparent” age, or an approximation of stand age based on measured variables other than time. A FSO old-growth classification is less subjective than current regression or indexing procedures, most of which assign an arbitrary value to each classification variable. In this example highlighting southern pine, five characteristic features of old-growth (q-factor, maximum tree DBH, stand basal area, percent red heart infection, and large woody debris volume) are expressed as response functions to help differentiate between stands synthesized from historical descriptions of virgin timber in southern Arkansas.

INTRODUCTION

Southern forests have been dramatically impacted by centuries of human influence. Logging, agriculture, settlement, introduced pests and pathogens, pollution, and other unnatural disturbances have notably altered natural ecosystems. Old-growth forests have been the most dramatically impacted: Davis (1996) estimated that less than 1% of original primary forests remain.

This rapid disappearance has placed modern-day public resource managers in a bind. Though expected to simultaneously protect threatened and endangered species (many of which are old-growth dependent), preserve and enhance existing resources, and maintain recreation and commodity production from a restricted land base, their options to address these issues are limited. New strategies like managing-for-old-growth have been proposed and are being implemented in small-scale field studies (e.g., Morton et al. 1991, Vora 1994), but the effectiveness of this approach has yet to be documented.

Management is further hindered by the lack of agreement on what constitutes old-growth (Hunter and White 1997). While the inherent differences between vegetation types precludes the development of a universal old-growth definition, there are also within-type issues. Different ecological thresholds such as minimum levels of woody debris or tree size are often considered, making it difficult to compare old-growth from one region (or study) to the next. Hunter and White (1997) noted that the arbitrariness of the current working definitions of old-growth did not improve management, as the thresholds used were

often based on limited criteria poorly related to stand potential.

Several researchers have cautioned against the single feature classification strategy (e.g., Franklin and Spies 1991, Rusterholz 1996, Hunter and White 1997), preferring instead an index of “old-growthedness” in which multiple factors are scored to produce an old-growth evaluation system. For example, Franklin and Spies (1991) proposed a continuous scaling strategy, thus allowing for various degrees of old-growthedness. By assuming these characteristics fit a “U” or “S” shaped curve, they employed an arbitrary scale to reflect different stand developmental stages. The resulting values could then be summed to produce an index of old-growthedness. Rusterholz (1996) described a similar approach that applied criteria based on cover type. Each candidate stand was given a value for each of these criteria based on predetermined thresholds, and then the cumulative score (to a maximum of 65 points) was used to determine its status. For example, pine forests were evaluated using the following criteria: stand age, size, and context; degree of human intervention; pine regeneration; tree size class diversity; maximum tree size; and large woody debris volume. Stands with scores of ≥ 40 were recommended for old-growth protection.

Another old-growth classification scheme was described by Hale et al. (1999), who applied a logistic regression model to differentiate between managed mature hardwood forests and unmanaged old-growth. They evaluated seven parameters before settling on large woody debris (LWD) volume as the most

¹ Research Forester, USDA Forest Service, Southern Research Station, P.O. Box 3516 UAM, Monticello, AR 71656. Phone: (870) 367-3464, Fax: (870) 367-1164, E-mail: dbragg@fs.fed.us

discriminating factor. While their fits were significant ($P < 0.05$), this procedure did not explain much of the variation in the data (adjusted $R^2 < 0.4$). This low explanatory power probably arose from the narrowness of old-growth defining criteria (LWD volume).

Part of the classification problem is the tendency to apply classical set theory when defining old-growth. In other words, a stand is either considered “old-growth” or “not old-growth,” usually with a singular threshold like stand age or average tree size. None of these systems have developed an objective approach in defining this critical threshold. Traditional set theory would require a threshold age (e.g., 200 years), for which anything older would be old-growth, and anything younger would not. But what do you do if a stand is 195 years old? A conceptually and mathematically rigorous multi-factor system to identify old-growth is needed. Even though numerical indexing (e.g., Franklin and Spies 1991, Rusterholz 1996) simultaneously incorporates multiple features in the identification of possible old-growth, it is prone to a high degree of subjectivity and statistical uncertainty. Logistic regression models are usually interpreted as all or nothing (even though their outcomes are fractional probabilities).

Recent developments in fuzzy set theory have provided a conceptual foundation that could address the problem of old-growth identification in a mathematically rigorous and ecologically sophisticated manner without burdening field managers with complex protocols and analysis procedures. Fuzzy sets, when generalized on dynamical systems theory (e.g., Roberts 1987a), produce a fuzzy systems theory which can then be used to determine the nature of the interaction between vegetative and environmental hyperspace (Roberts 1989). Thus, not only does fuzzy mathematics provide a more intuitive approach to many ecological questions, but it incorporates the dynamics involved between the key components of ecological systems.

A fuzzy set ordination (FSO) approach to old-growth classification represents an obvious departure from traditional approaches to old-growth delineation. The development of FSO old-growth classification was predicated on the following principles: 1) it must improve upon current classification procedures, and 2) it should be easily applied using traditional field measurements. This paper outlines conceptual and mathematical principles for a fuzzy set classification of old-growth southern pine.

FUZZY SET ORDINATION

One of the major advantages of fuzzy sets is that they preserve the algebra of set theory, thus retaining the formal logic and mathematics of “crisp” (Euclidian) sets. Fuzzy mathematics can work on either continuous or discrete variables. FSO has considerable

potential in ecological analysis because it does not depend on specific thresholds, but rather membership in fuzzy sets based on their degree of similarity to a reference set (Roberts 1989). Fuzzy set theory scores attributes based on their similarity to the largest (or upper) limit considered and assigns them a membership in the intended classification set. Thus, a stand that is 195 years old would receive a very high membership (probably > 0.95) in the fuzzy set of old stands. Traditional set theory, conversely, would reject this a stand as old-growth because it did not meet the minimum age threshold.

The following sections provide a brief synopsis of fuzzy set mathematics (more detailed reviews can be found in Roberts (1987b, 1989)). All parameters in this study were indexed to range from 0 to 1, with 1 representing the maximum value of that variable and 0 indicating the lowest value. As an example, a linear indexing would follow:

$$V = \frac{V_{actual} - V_{min}}{V_{max} - V_{min}} \quad (1)$$

where the indexed value (V) is a function of the actual (V_{actual}), minimum (V_{min}), and maximum (V_{max}) value of that parameter. By scaling variables between 0 and 1, the data became self-calibrated so that specific thresholds did not have to be predetermined. In traditional crisp sets, a stand would have membership in each of the defined sets as a 0 (absent) or 1 (present). Fuzzy sets allow for fractional membership in sets such that a stand could be anywhere in the range from 0 to 1. Each corresponding parameter set (generically referred to as P) are denoted by italicized capital letters. Thus,

$$P = \{(x, \mathbf{m}_p(x))\} \quad (2)$$

where x is an element of P and $\mathbf{m}_p(x)$ is the membership of x in set P .

Because fuzzy sets retain the mathematics of traditional set theory, the following operators are defined for generic sets P and Q :

union–

$$\mathbf{m}_{p \cup q}(x) = \max(\mathbf{m}_p(x), \mathbf{m}_q(x)) \quad (3)$$

complement–

$$\mathbf{m}_{\bar{p}}(x) = 1 - \mathbf{m}_p(x) \quad (4)$$

difference–

$$\mathbf{m}_{p - q}(x) = \min(\mathbf{m}_p(x), \mathbf{m}_{\bar{q}}(x)) \quad (5)$$

intersection–

$$\mathbf{m}_{p \cap q}(x) = \min(\mathbf{m}_p(x), \mathbf{m}_q(x)) \quad (6)$$

Union, intersection, and complement can be considered “and,” “or,” and “not,” respectively

(Roberts 1986). Thus, in a multi-factor fuzzy set classification, one is first interested in how the sets orient themselves and then on defining factors. Roberts (1987b) adapted these operators to produce a new approach to sets: the anticommutative difference operator (ADO). The ADO (which can be considered as “while not”) allows for contrasts between dissimilar sets:

$$m_{p \vee q}(x) = \frac{[1 + (m_q(x))^2 - (m_p(x))^2]}{2} \quad (7)$$

The ADO allows for complex sets based on gradients to be developed. Since most environmental gradients tend to be complementary, stands can be considered “similar” to one end while not similar to the other. To arrive at this stage, plot variables must be associated with a similarity index. While many such indices abound, I applied Roberts’ Index (Roberts 1986):

$$S_{xy} = \frac{\sum_{i=1}^n \left[(V_{xi} + V_{yi}) \left(\frac{\min(V_{xi}, V_{yi})}{\max(V_{xi}, V_{yi})} \right) \right]}{\sum_{i=1}^n (V_{xi} + V_{yi})} \quad (8)$$

where S_{xy} is the similarity of stand x to stand y , and V is the indexed value calculated earlier for each of n parameters. It is within the calculation of S_{xy} that the parameters are combined to allow for a multi-factor old-growth classification. Since the primary consideration in determining old-growth is the age of the stand, I shall frame this in terms of age and “apparent” age. Age is directly measured for each stand, while apparent age is the fuzzy prediction deduced from their similarity to either young or old stands. To avoid circular conflicts, I did not use age as one of the indexed values. FSO is the juxtapositioning of apparent age with actual age. The set of stands *similar* to old stands (set O) was then:

$$m_o(x) = \frac{\sum_{y \neq x} [S_{xy}(m_A(y))]}{\sum_{y \neq x} (m_A(y))} \quad (9)$$

where $o(x)$ is the membership of stand x in set O and $A(y)$ is the membership of stand y in set A (in this case, the set of old stands). Note that membership in O is related to similarity to stands in set A using equation (9). Likewise, the set of stands similar to young stands (set Y) is:

$$m_y(x) = \frac{\sum_{y \neq x} [S_{xy}(m_Y(y))]}{\sum_{y \neq x} (m_Y(y))} \quad (10)$$

where $y(x)$ is the membership of stand x in set Y and $o(y)$ is the membership of stand y in set O (young stands, or the complement of the set of old stands). The newly defined sets ($o(x)$ and $y(x)$) can then be placed in the ADO equation, re-standardized to range between 0 and 1, and then compared to measured stand age to indicate their position along the sere.

FSO results can be interpreted in several ways. Scores from individual stands can be ranked and evaluated. For example, it would be possible to use the apparent age gradient as a scaling for old-growthedness. Obviously, a score = 1 would indicate full membership in the set of old-growth stands, suggesting that all measured parameters were optimally met by this case. When scores fall between 0 and 1, then some condition(s) are less than maximum for a stand of a given age, which may or may not preclude the stand from further consideration as old-growth. Minimum levels of old-growthedness based on desired conditions could then be identified and managed for. For instance, old-growth reserve (i.e., no treatment) stands may have a value of 0.8 or greater, while those ranging from 0.6 to 0.8 could be considered as candidates for specialized treatment. Fuzzy set ordination scores could also be used to evaluate residual differences from the ordination graph and hence prove useful in identifying deficient or excessive conditions.

METHODS

Using a set of derived gradients based on synthetic (but ecologically reasonable) trends for southern pine stands, a fuzzy set ordination was performed to anticipate stand age solely as a function of these parameters.

Cover type selection and period delineation

The first step in any old-growth classification is the identification of the relevant cover type and time period. This is critical because one would not expect the parameters of interest for old-growth loblolly (*Pinus taeda* L.) and shortleaf pine (*Pinus echinata* Mill.) stands to be the same as those for baldcypress (*Taxodium distichum* (L.) Rich.) stands. The desired time period should also be identified, as conditions may also vary temporally. This effort considered factors for the virgin loblolly and shortleaf pine-dominated ecosystems of the Upper West Gulf Coastal Plain of Arkansas during the early 19th Century because 1) these forests were once common, but now are very limited; 2) they have an existing historical and contemporary literature base from which to parameterize; and 3) there is on-going research into managing-for-old-growth conditions, thus supporting the development of evaluative criteria.

Parameter selection

Any parameter with a functional relationship to stand age could be used (Roberts 1986). Conditions that were specifically quantifiable and unambiguous in the literature on old-growth pine were selected (rather than vague concepts like “absence of human disturbance”). Because the intention of this paper is to generally illustrate the FSO classification strategy, the values

Figure 1. Synthetic trends of parameters used to define response curves for fuzzy set ordination of old-growth southern pine (also see Appendix A).

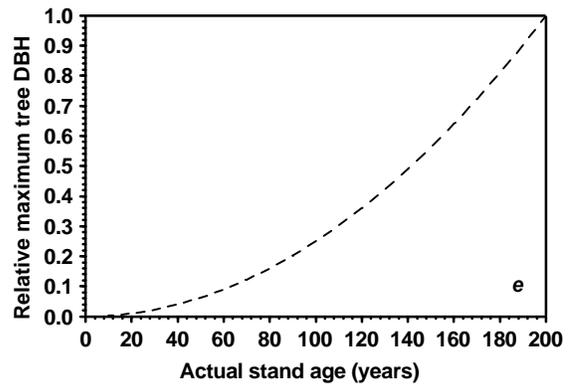
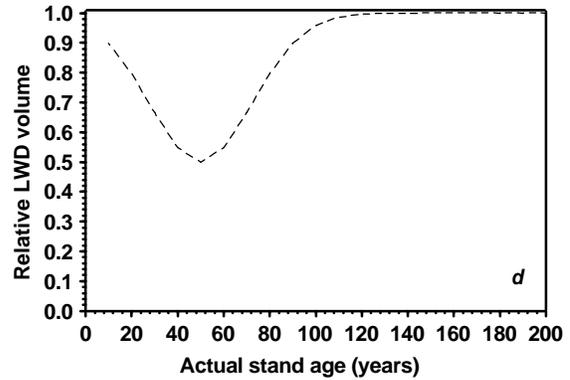
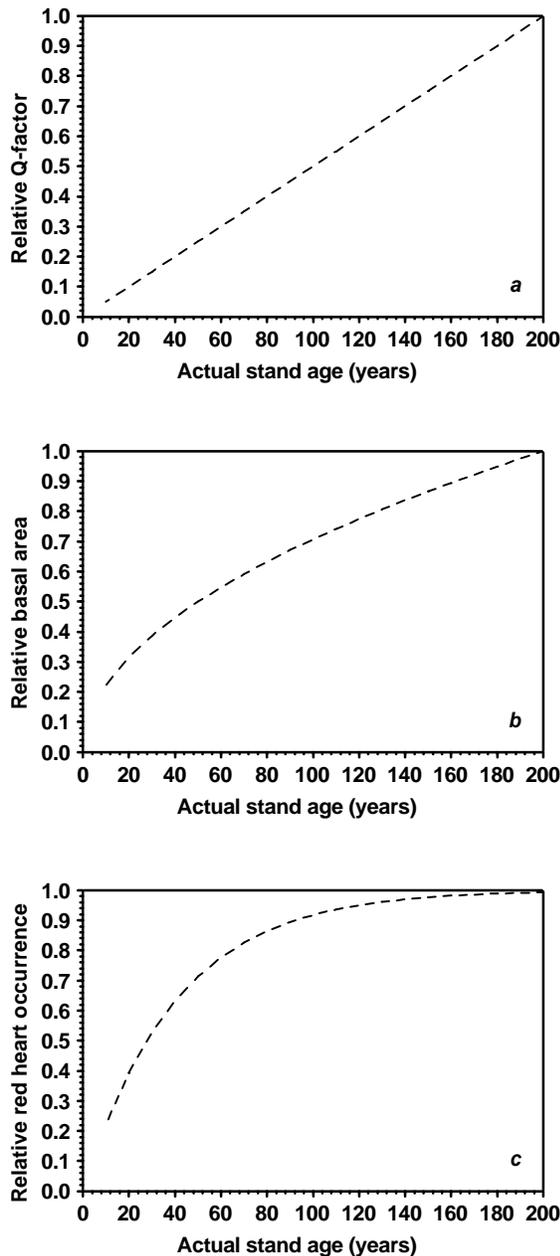


Figure 1 (cont.). Synthetic parameter trends.

presented were synthesized (without variance) from reasonable trends (Appendix A). Twenty stands were assembled from these synthetic values (Figure 1), combined using the Roberts similarity index, and then processed to produce an interpretable FSO.

The attributes used in this analysis included q factor, maximum tree diameter at breast height (DBH), stand basal area, red heart (*Phellinus pini* Ames) abundance, and LWD volume. These features are primarily structural, but should be well correlated with other less tangible old-growth attributes. Q factor is an abstraction of the relationship between stocking and diameter class, with higher numbers indicating a steeper trend (more small trees, few large ones) and a lower number suggesting fewer small trees and more big ones (typical of old-growth) (Smith 1986). Maximum tree DBH indicates the upper end of the structural condition of the forest, while stand basal area integrates size and stocking to suggest developmental stage. Red heart is a fungal heart rot that increases markedly as pine ages (Mattoon 1915). Dead wood volume is also strongly suggestive of developmental stage: old-growth usually contains substantial quantities of large LWD, while managed stands do not (e.g., Gore and Patterson 1986, Goodburn and Lorimer 1998, Hale et al. 1999).

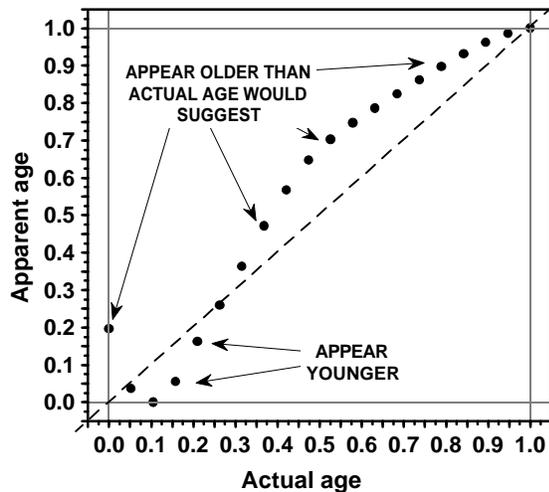


Figure 2. Apparent age (predicted from the fuzzy set ordination) compared to actual age (200 years when actual age = 1.0).

RESULTS

Fuzzy set ordination did a good job of predicting stand age from the variables it was provided (Figure 2). In general, the younger stands had attributes less like old-growth, while old stands were quite similar.

Residual differences are the deviations from the equivalence (dashed) line in Figure 2, and can be either positive or negative. Some stands appeared older than their chronological age would otherwise indicate, while others appeared younger than expected. The obvious departures from the 1:1 line in Figure 2 can be best understood by considering the features most responsible for this behavior. The bowing of the ordination results in Figure 2 is associated primarily with red heart abundance (Figure 1c). With the assumption of this study, the stands are noticeably overstocked with heart rot from a 1:1 expectation. The deviation apparent in young stands arose from the higher-than-expected volume of LWD present in these stands (Figure 1d). Large quantities of LWD are not unusual in young stands, especially those arising after catastrophic natural disturbances or timber harvesting (Sturtevant et al. 1997).

DISCUSSION

FSO versus numerical indexing

While the Franklin and Spies (1991) and Rusterholz (1996) procedures are more holistic than simple thresholds, they contain considerable subjectivity in their determination of old-growth point values. Since there is no mathematical basis to the values assigned, it could be argued that other sets of features or different emphasis on the criteria may result in a dramatically dissimilar outcome. FSO ordination avoids this issue because the measurements are scaled to those found in stands indisputably considered old-

growth.

FSO versus logistic regression

A fuzzy set approach to old-growth classification is also an improvement over logistic regression analysis. Perhaps the biggest problem with a logistic approach is that it is inherently circular: to fit the regression, a stand must be classified *a priori* as “old-growth” or “not old-growth,” and then the coefficients are determined. Thus, using the resulting probability to categorize old-growth would not yield independent predictions. FSO does not require a defining variable like actual age to predict apparent age, and therefore avoids the problem of circularity. Additionally, the fitted nature of multivariate regression limits the interpretability of the residuals, and thus provide less utility in using that system to adaptively manage old-growth.

Potential applications

The interpretation and management directions suggested by residual analysis are some of the prime advantages to FSO. Identifying the factors leading to these discrepancies could be directly used to manage particular areas considered old-growth. Perhaps a stand appears younger than expected because of unusually low levels of LWD. This deficiency could be accommodated by the creation of new snags and/or downed logs. Individual parameters could be tested for their relative importance on the fuzzy old-growth classification by simple correlation analysis. Noticeable patterns may arise over part or all of the age gradient, which in turn can lead to further management emphasis on those components most sensitive to the correlation analysis.

The flexibility permitted by not having to define old-growth criteria *a priori* should also allow better customization of the process. This method also lacks the subjectivity of previous indexing methods as each variable used in the final analysis has been self-calibrated (as opposed to arbitrarily scored). The ability to combine multiple factors in an objective process will also improve classification from systems that key upon a single factor.

Limitations and pitfalls of the method

The success of a fuzzy approach to old-growth classification depends on our ability to identify clear patterns between stand age and parameters assumed to be indicative of old-growth-like conditions. Since old-growth stands are notoriously variable, only poor trends may appear, resulting in a weakly correlated classification outcome. FSO, however, is surprisingly robust to noise (Roberts 1998), so weak trends (noisy data) are not as detrimental to FSO as with other statistical approaches. It is also vital to sample a reasonably long temporal developmental gradient to

help identify the key factors for classification because disturbances may cloud some of the relationships between stand structure and age (e.g., storm-related LWD accumulation).

CONCLUSIONS

Fuzzy set ordination appears to have considerable promise for old-growth classification. Even with a limited amount of structural parameters, it was possible to recover most of the structure of a synthetic gradient of different aged stands without specifically using age to organize the stands. FSO permits the direct interpretation of deviations from expected values in a manner rarely available for most old-growth classification strategies. This in turn suggests that management activities could be planned from the outcome of the ordination to optimize the value of existing stands for future action.

ACKNOWLEDGMENTS

I would like to thank Dave Roberts (Utah State University) for his inspiration and support of this effort, especially during its early stages. Thoughtful review comments by Mike Shelton (USDA Forest Service) and Hope Bragg (University of Arkansas-Monticello) improved the quality of this paper.

LITERATURE CITED

- Davis, M.B. 1996. Extent and location. *In* Eastern Old-Growth Forests: Prospects for Rediscovery and Recovery, M.B. Davis, ed. Island Press, Washington, DC. p. 18-32.
- Franklin, J.F. and T.A. Spies. 1991. Ecological definitions of old-growth Douglas-fir forests. *In* Wildlife and Vegetation of Unmanaged Douglas-fir Forests, L.F. Ruggiero, K.B. Aubry, A.B. Carey, and M.H. Huff, tech. coord. USDA Forest Service General Technical Report PNW-285. p. 61-69.
- Goodburn, J.M. and C.G. Lorimer. 1998. Cavity trees and LWD in old-growth and managed northern hardwood forests in Wisconsin and Michigan. *Canadian Journal of Forest Research* 28:427-438.
- Gore, J.A. and W.A. Patterson. 1986. Mass of downed wood in northern hardwood forests in New Hampshire: potential effects of forest management. *Canadian Journal of Forest Research* 16:335-339.
- Hale, C.M., J. Pastor, and K.A. Rusterholz. 1999. Comparison of structural and compositional characteristics in old-growth and mature, managed hardwood forests of Minnesota, U.S.A. *Canadian Journal of Forest Research* 29:1479-1489.
- Hunter, M.L. and A.S. White. 1997. Ecological thresholds and the definition of old-growth forest stands. *Natural Areas Journal* 17:292-296.
- Mattoon, W.R. 1915. Life history of shortleaf pine. *USDA Bulletin* 244. 46 p.
- Morton, D. and the New Perspectives Team. 1991. Maintaining old growth in northern hardwood ecosystems. *In* Proceedings of the 1991 Society of American Foresters National Convention, SAF Publication Number 91-05. p. 527-528.
- Roberts, D.W. 1986. Ordination on the basis of fuzzy set theory. *Vegetatio* 66:123-131.
- Roberts, D.W. 1987a. A dynamical systems perspective on vegetation theory. *Vegetatio* 69:27-33.
- Roberts, D.W. 1987b. An anticommutative difference operator for fuzzy sets and relations. *Fuzzy Sets and Systems* 21:35-42.
- Roberts, D.W. 1989. Fuzzy systems vegetation theory. *Vegetatio* 83:71-80.
- Roberts, D.W. 1998. Fuzzy set ordinations are resistant to noise and inappropriate transformations of environmental data. Oral Presentation at the 83rd Annual Meeting of the Ecological Society of America. *Ecological Society of America Abstracts*, p. 112.
- Rusterholz, K.A. 1996. Identification and protection of old growth on state-owned land in Minnesota. *In* Eastern Old-Growth Forests: Prospects for Rediscovery and Recovery, M.B. Davis, ed. Island Press, Washington, DC. p. 233-244.
- Smith, D.M. 1986. The practice of silviculture. 8th ed. John Wiley & Sons, New York, NY. 527 p.
- Sturtevant, B.R., J.A. Bissonette, J.N. Long, and D.W. Roberts. 1997. Coarse woody debris as a function of age, stand structure, and disturbance in boreal Newfoundland. *Ecological Applications* 7:702-712.
- Vora, R.S. 1994. Integrating old-growth forest into managed landscapes: a northern Great Lakes perspective. *Natural Areas Journal* 14:113-123.

Appendix A. Realization of synthetic^a trends assumed in Figure 1, including both actual and indexed (Rel.) values.

Stand Number	Stand Age (yrs.)	Q Factor (unitless)	Rel. Q Factor	Max. DBH (cm)	Rel. Max. DBH	Basal Area (m ² /ha)	Rel. Basal Area	Red Heart (%)	Rel. Red Heart	LWD Volume (m ³ /ha)	Rel. LWD Volume
1	200	1.1	1.0000	140.0	1.0000	20.0	1.0000	19.9	0.9933	40.0	1.0000
2	190	1.2	0.9500	127.3	0.9025	20.5	0.9747	19.8	0.9913	40.0	1.0000
3	180	1.3	0.9000	115.3	0.8100	21.0	0.9487	19.8	0.9889	40.0	1.0000
4	170	1.4	0.8500	103.9	0.7225	21.6	0.9220	19.7	0.9857	40.0	1.0000
5	160	1.5	0.8000	93.2	0.6400	22.1	0.8944	19.6	0.9817	40.0	1.0000
6	150	1.6	0.7500	83.1	0.5625	22.7	0.8660	19.5	0.9765	40.0	1.0000
7	140	1.7	0.7000	73.7	0.4900	23.3	0.8367	19.4	0.9698	40.0	0.9998
8	130	1.8	0.6500	64.9	0.4225	23.9	0.8062	19.2	0.9612	40.0	0.9992
9	120	1.9	0.6000	56.8	0.3600	24.5	0.7746	19.0	0.9502	39.9	0.9963
10	110	2.0	0.5500	49.3	0.3025	25.2	0.7416	18.7	0.9361	39.5	0.9864
11	100	2.1	0.5000	42.5	0.2500	25.9	0.7071	18.4	0.9179	38.4	0.9590
12	90	2.2	0.4500	36.3	0.2025	26.6	0.6708	17.9	0.8946	36.0	0.8991
13	80	2.3	0.4000	30.8	0.1600	27.4	0.6325	17.3	0.8647	31.9	0.7968
14	70	2.4	0.3500	25.9	0.1225	28.2	0.5916	16.5	0.8262	26.6	0.6649
15	60	2.5	0.3000	21.7	0.0900	29.0	0.5477	15.5	0.7769	21.9	0.5476
16	50	2.6	0.2500	18.1	0.0625	30.0	0.5000	14.3	0.7135	20.0	0.5000
17	40	2.7	0.2000	15.2	0.0400	31.1	0.4472	12.6	0.6321	21.9	0.5476
18	30	2.8	0.1500	12.9	0.0225	32.3	0.3873	10.6	0.5276	26.6	0.6649
19	20	2.9	0.1000	11.3	0.0100	33.7	0.3162	7.9	0.3935	31.9	0.7968
20	10	3.0	0.0500	10.3	0.0025	35.5	0.2236	4.4	0.2212	36.0	0.8991

^a Trends are “synthetic” in that they reflect reasonable estimates of a parameter at the given age of the stand, but do not represent field-measured values. Due to rounding, some actual values may not precisely correspond to indexed ones.