Modeling / Inventory

USING DYNAMIC PROGRAMMING TO EXPLORE HARDWOOD SILVICULTURAL REGIMES

Matthew H. Pelkki[1]

Abstract—Forest managers face a difficult problem when required to estimate future physical and economic returns from timber stands when management strategies can vary by number and timing of thinnings, thinning type, thinning intensity and final harvest age. Field-based testing of these variables provides the most reliable information on stand and individual tree response to these practices, but is limited by the time and cost of these studies. Simulation studies based on growth models which derive their tree and stand response data on field observations are one method of exploring the response curve in greater detail and at a much lower cost in time and effort. Dynamic programming is an ideal method for exploring both stand and individual-tree response to new economic conditions and rotation schemes. Dynamic programming can be used to investigate possible responses to different thinning strategies, intensities, and timing of thinnings. It can also provide insights into management responses to changing markets and prices for wood products. However, simulation approaches are hindered by the weakness of long term (40+ years) growth projections and the ability of the computer systems to accurately model real-world constraints. In this paper, a dynamic programming (DP) based computer optimization technique is used and compared with the results of DP-based studies to previously published field studies of yellow-poplar. The similarities found lend credence to the biological reasonableness of the simulations and the ability of the simulation to provide detailed sensitivity analysis demonstrates its strengths in identifying new research questions for field-based studies.

INTRODUCTION — TIMBER MANAGEMENT

Timber managers face a myriad of questions in planning a stand from regeneration to final harvest. Two of the most basic questions center around rotation length and stand density. These two questions are appropriate whether the objective is fiber or financial maximization. To answer these questions, foresters must pull out the “crystal ball” and make predictions, using some form of growth and yield model to estimate future stand characteristics (total volume, diameter distribution, species composition). If financial maximization is desired, then foresters must also attempt to predict future economic conditions such as the cost of capital, inflation, future stumpage prices, and the costs of harvesting. These data are important for making decisions about sustainable forest-wide management decisions as well as good stand-level decisions. For the forest cover types in the central United States, this usually involves planning and prediction of forest and financial conditions on a time frame from 25 to 100 years.

Two approaches have been applied in an effort to provide forest managers with some guidelines for decision making about rotation length and stand density. The first and oldest technique is field testing. Representative sites are selected and management alternatives are actually implemented and the physical results are measured and reported. The second approach has grown in use since the advent of powerful, low-cost computers. Computer simulation of a forest stand is based upon the use of growth and yield models and this is coupled with economics to provide estimations of future outputs.

The following discussion will compare these two approaches and then provide a detailed description and one example of an implementation of dynamic programming which represents one technique for computer simulation of forest stands. One overall assumption will be that the problem under consideration is an even-aged stand, although all techniques are equally applicable to uneven-aged stand structures.

TWO APPROACHES — COMPUTER SIMULATION AND FIELD TESTING

Necessary Resources for Each Method

Actual field testing of silvicultural regimes requires access to representative sites and the ability to actually implement and then observe the results of management actions. Considerable financial, personnel and equipment resources are needed. This method also requires time for the stand to respond to the management actions and for repeated monitoring of stand parameters. It is assumed that environmental conditions (site, climate) will be typical of the region to which the study applies.

Computer simulation, on the other hand, requires good data on initial stand conditions as a starting point for the simulation. These initial conditions may be a selection of actual representative sites or a hypothetical composite site that is considered typical of the region. Growth and yield models that are responsive to the range of actions and stand conditions simulated are necessary. For example, the TWIGS individual-tree growth simulator for the Central United States (Miner and others 1988) was developed from a data base that poorly represented yellow-poplar, Eastern redcedar, and tupelo/gum. Therefore, this growth simulation would not be an optimal choice for computer simulation of these species in the Central United States. Also, whole stand models such as

[1] Associate Professor of Forest Management and Economics, University of Kentucky, Department of Forestry, Lexington, KY 40546-0073.

Citation for proceedings: Stringer, Jeffrey W.; Loftis, David L., eds. 1999. Proceedings, 12th central hardwood forest conference; 1999 February 28-March 1-2; Lexington, KY. Gen. Tech. Rep. SRS-24. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 293 p. [Peer-reviewed paper].

yield tables are unable to accurately predict how a stand might grow if the diameter distribution is altered from thinnings from above or below. Predicting future tree quality is also very difficult with whole stand models. Furthermore, computer simulations will need estimates of future financial conditions. Data or additional models that predict stumpage prices, the cost of capital (interest rates), and the cost of harvesting operations must be included.

Strengths and Weaknesses of Each Approach

Table 1 is a list of some major strengths and weaknesses of each approach. Field-based studies have several advantages. They do not rely upon predictions of growth and yield or economic conditions. It is difficult to argue with observed results when properly replicated. However, it should be noted that even observed economic results, when applied in planning future rotations, assume that the future will be the same as the past. The alternatives for management that are chosen can be put into practice with relative ease. Due to the great length of time required and the costs related to acquiring the sites and managing them, field-based studies are expensive. Sensitivity analysis is also difficult and type and number of alternatives is usually kept very low so that adequate replication can be achieved.

Computer simulation, on the other hand, can test a great number of alternatives and combinations. Due to the speed and power of modern personal computers, simulations can test millions of combinations in hours rather than decades and at a far lower cost. However, long-term projections with growth and yield models may be substantially in error (Rauscher and others 1997) and at best might be considered only relatively correct. Another subtle, but important problem with computer simulation is "artifacts." An artifact represents a difference in a result that is not likely to be observed in the real world. For example, Beck and Della-Bianca developed yield equations for unthinned (1970) and thinned (1975) yellow-poplar stands. The two equations are disjoint at any single given age. When these two models are used together in a computer simulation, when a stand is thinned, it immediately shows a slight increase in volume at the same age! This directs the simulation to thin at an early age to capture this "free" volume which, in reality, does not exist. Another form of an artifact may occur from rounding of values in a simulation. For example, it may be possible to test the sensitivity of thinning in basal area increments of one square foot or less, but would the results be practicable in the real world?

The remainder of this paper presents the basics of one computer simulation technique, dynamic programming (DP). A non-mathematical description of DP will be presented, followed by a mathematical formulation and then a description of a DP-based computer program (NESTER). The results of published research on field-based studies of yellow-poplar will be compared to some NESTER results.

OPTIMIZATION USING DYNAMIC PROGRAMMING

Dynamic programming is a computational method that takes large sequential problems and breaks them down into solvable, related subproblems which can be linked together to achieve an optimal solution to the entire problem. Let's think about a 20-year-old even-aged hardwood stand. Let us further assume that management operations will occur only once every 10 years, and that the alternatives available are, to do nothing, thin 20 percent or 40 percent of the volume, or clear fell the entire stand (regeneration harvest). Figure 1 shows a diagram with a node or state representing our 20-year-old stand. The state is described by the total cubic foot volume per acre, in this case 700 ft^3/acre. Over the 10 years, this stand will grow 700 ft^3/acre and will have a total of 1400 ft^3/acre. If we do nothing, the stand remains at 1400 ft^3/acre. If we thin, the stand volume will be reduced to either 1120 ft^3/acre (20 percent thinning) or 840 ft^3/acre (40 percent thinning). If we clear fell the stand it has no residual volume. From one state in the initial stage (age 20), we now have four possible states in the first stage. Figure 2 shows a partial diagram of the progression into the second stage. Note that two paths compete for the state described by 1600 ft^3/acre. As each stage continues, we will have more states in the decision space and as that happens, it is more likely that two or more paths will

Table 1—Strengths (+) and weaknesses (–) of field-based studies and computer simulations

Field-based	Computer Simulation
+ Long term growth and yield valid	– Long term problems with growth and yield models
+ Economic data is observed, not predicted	– Long term economic assumptions can be invalid
+ Alternatives selected are relatively easy to put into practice	– Model may reflect artifacts based on structure of model
– Costly to implement	+ Inexpensive to test
– Difficult to test large numbers of alternatives and combinations/intensities	+ Excellent sensitivity analysis
– Long time for study completion	+ Fast results

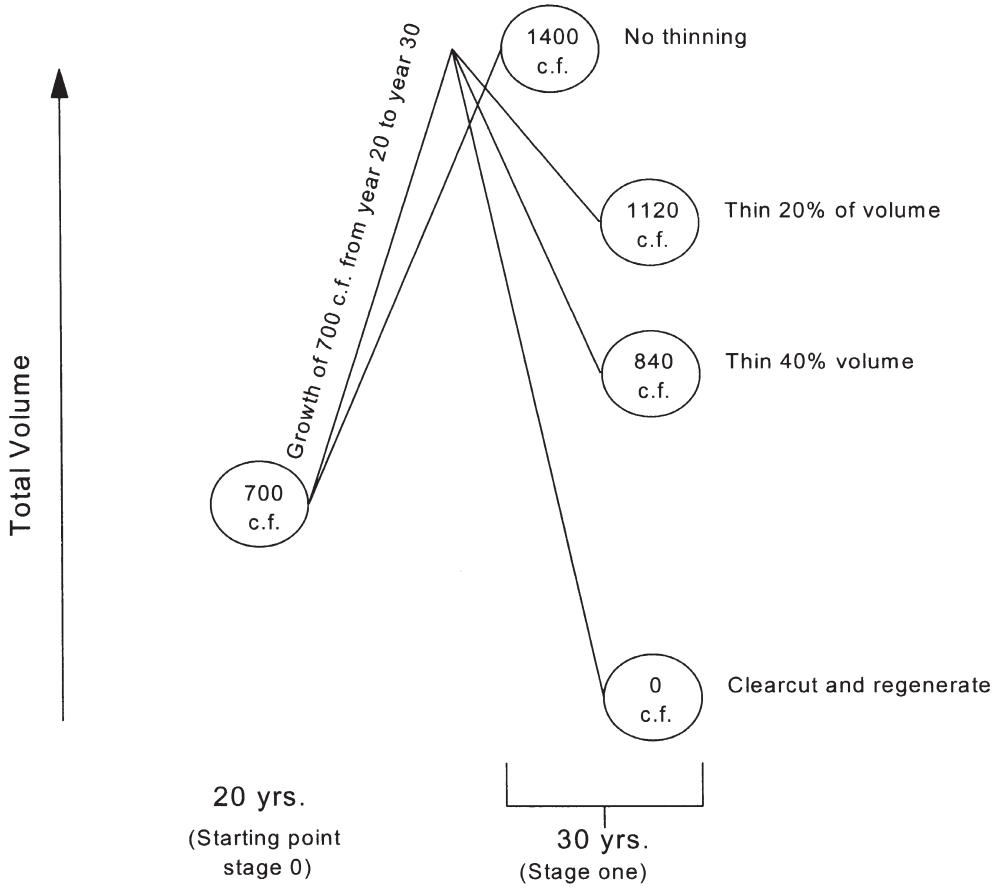

Figure 1—Initial condition and stage one states for forward-recursive dynamic programming network.

compete for a given state. The underlying assumption in dynamic programming is that any two or more paths that compete for the same state represent an identical forest condition with the same future value growth potential. If we assume that it doesn't matter how we achieve 1600 ft^3/acre in a 30-year-old stand, this assumption holds true. This is called the principle of optimality (Dykstra 1984). It is also called the memoryless principle and it means that an optimal path through the network of all possible states can be solved one stage at a time.

More formally, the subproblems are often referred to as stages, and within each stage are multiple states. A recursive relationship links the stages together. This recursive relationship contains a method for choosing optimal states within a stage and for linking optimal stage policies together into an optimal problem solution. To do so, certain information must be stored at each state. This includes 1) a link to a state in an adjacent stage, 2) the objective function value for reaching this particular state, and finally, 3) information necessary to describe the state and make a decision regarding options for moving it to the next adjacent state and determining its value relative to the objective function (Figure 3). The formal mathematics for dynamic programming are provided in the next section.

Dynamic Programming Formulated Mathematically

The objective function shown in equation one relates the value of N management decisions, $f_N^*(Y_N)$, which yield a stand described by Y_N (a regeneration harvest in even-aged stands). A management decision is defined by the variable T_n.

$$f_N(Y_N) = \sum_{n=0}^{N} r_n(T_n) \qquad (1)$$

Equation two relates the current state of a stand in stage n to a state in the next stage n+1 ($Y_n \rightarrow Y_{n+1}$) by taking the current stand condition (Yn) and adding the next current growth ($G_{n+1}(Y_n)$) and subtracting out any intermediate harvests (T_{n+1}).

$$Y_n + G_{n+1}(Y_n) - T_{n+1} = Y_{n+1} \qquad (n=0,1,2,\ldots,N-1) \qquad (2)$$

Equation three simply relates the volume of a stand after it has grown (X_n) to its final volume in stage n (Y_n) by showing the difference as the amount of volume harvested (T_n).

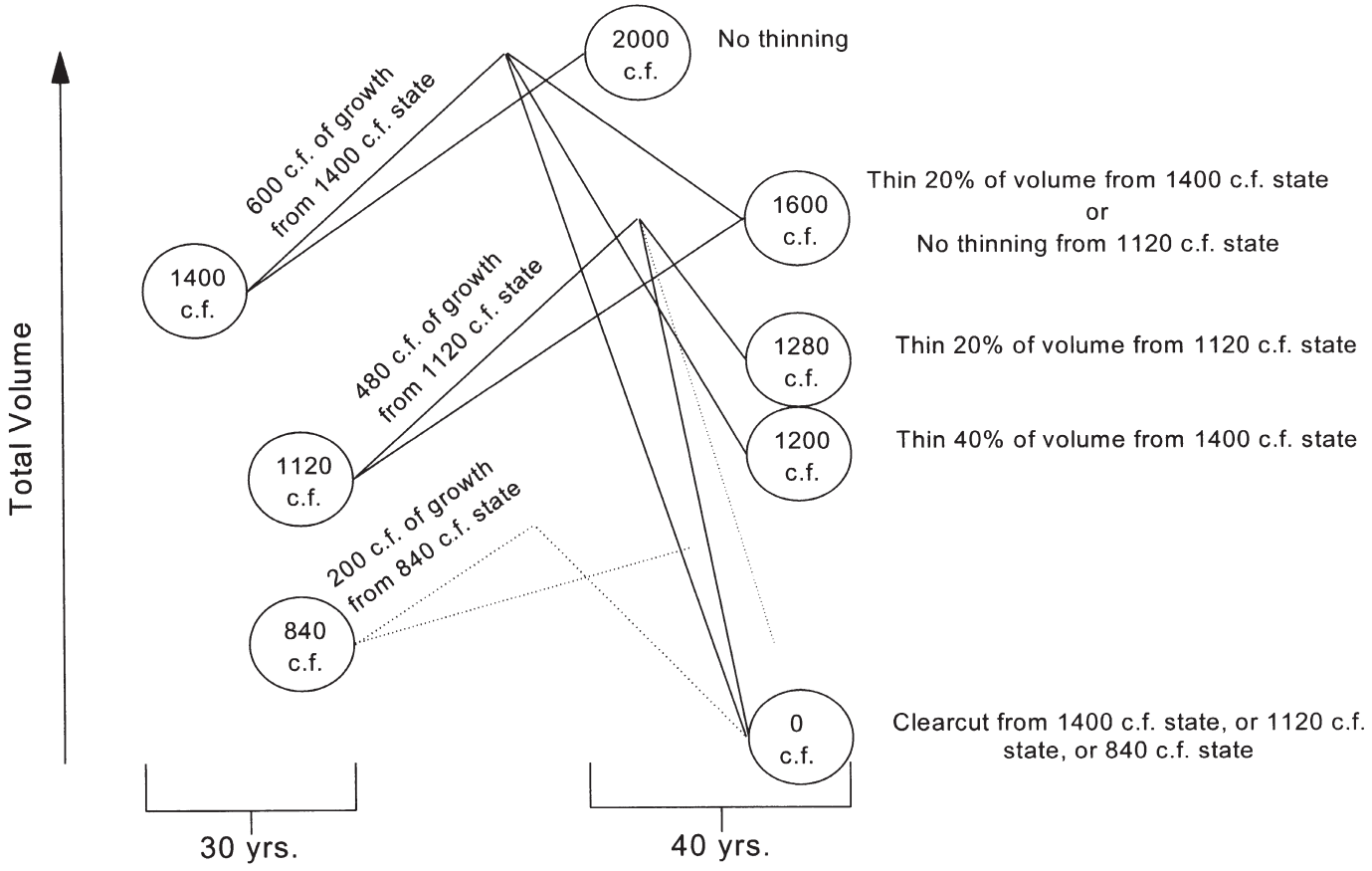

Figure 2—Stages one and two of a dynamic programming network.

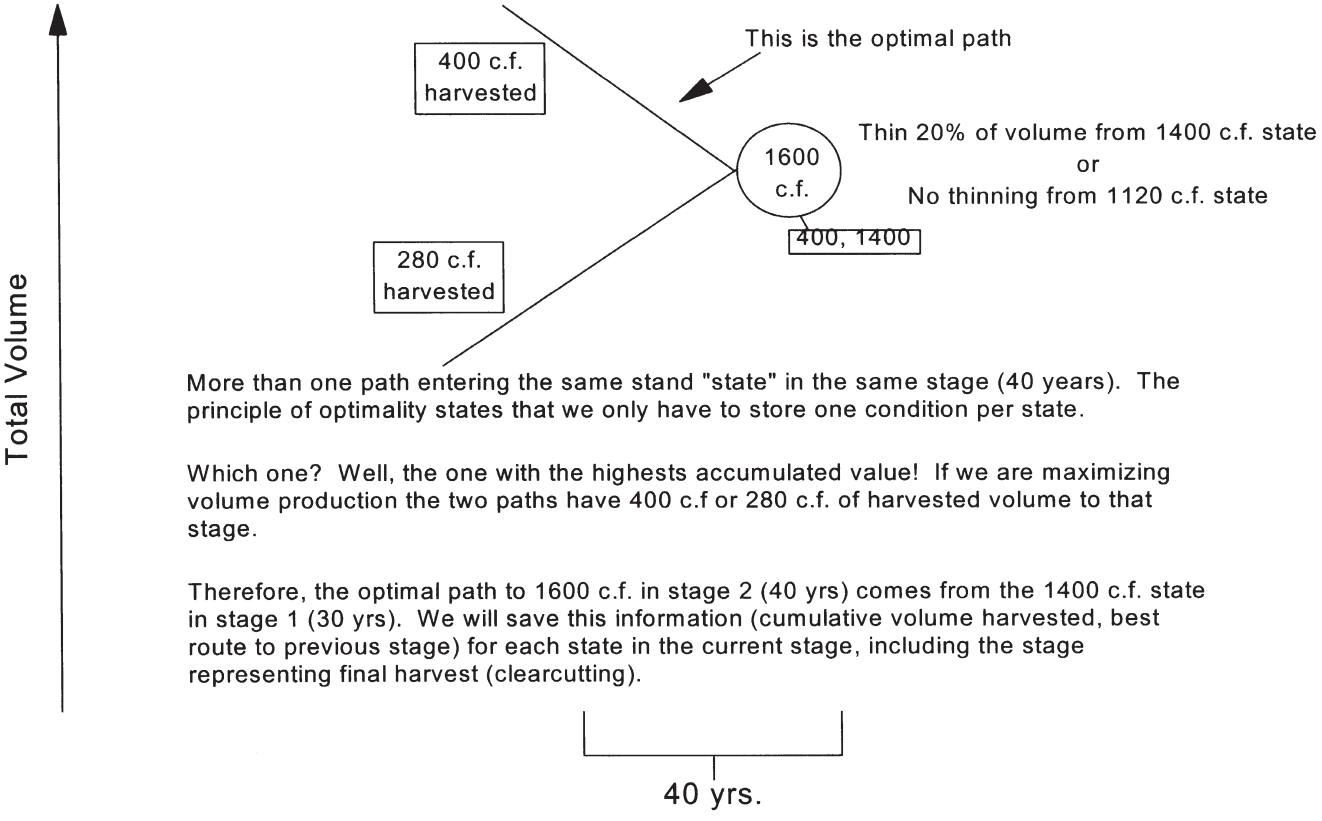

Figure 3—Competition for nodes in the DP network and tracing the optimal path.

$$X_n - T_n = Y_n \qquad (n=0,1,2,\ldots,N) \qquad (3)$$

Equation four assures that the final stand condition is a clearcut. If uneven-aged management is applied, then a different ending condition, one with a residual volume, is defined.

$$X_N - T_N = 0 \qquad (4)$$

Equation five is the recursive relationship that links each stage together. The function $r_n(X_n,T_n)$ is the return generated at stage n from decision T_n taken on stand described by X_n. The function r can maximize any quantifiable criteria but is usually a maximum volume or value function.

$$f_n^*(Y_n) = \max_{(Y_{n-1},T_n)} [r_n(X_n,T_n) + f_{n-1}^*(Y_{n-1})] \qquad (5)$$

The variables are defined again for the reader:

$f_N^*(Y_N)$ = objective function value of N management decisions yielding a stand described by Y_N. In the final stage, N, Y_N is a clear cut stand.

r_nT_n = return generated at stage n by decision T_n.

X_n = state vector describing stand at stage n after it grows from state Y_{n-1}.

Y_n = state vector describing stand at stage n.

T_n = management decision taken at stage n.

$G_{n+1}(Y_n)$ = growth of stand at stage n to stage n+1. Along with T_n, this constitutes the transformation function.

$r_n(X_n,T_n)$ = return generated at stage n from decision T_n taken on stand described by X_n.

Some Comments About Dynamic Programming

A single variable such as total volume is usually insufficient to adequately describe a stand. Two stands, both having a total volume of 1000 ft^3/acre may have very different average diameters and thus different future growth and value potential. Therefore, two or more state variables are used to define a stand condition within a stage. However, this leads to another problem known as the curse of dimensionality (Arthaud 1986). State variables used to describe forest stands are usually continuous. If the variables cubic volume per acre and number of trees per acre are used to define states within a stage and it is assumed that the maximum volume is 5000 ft^3/acre and the maximum number of trees is 500/acre, then the DP program must be able to represent 2.5 million states per stage (assuming that all volumes and trees per acre are rounded to the nearest integer). If yet another state variable is added, the memory requirements are likely to exceed the primary memory capacity of most University mainframe computers. To effectively use continuous variables and to reduce the size of the solution network, the idea of state neighborhoods (Brodie and Kao 1979) was introduced. For example, state neighborhoods of ± 5 trees per acre and ± 5 ft^3/acre would reduce the two state memory requirements from 2.5 million to 100,000 states per stage. Figure 2 demonstrates an example of state neighborhoods. If the neighborhood interval was 100 ft^3/acre, then the paths leading to 1280 ft^3/acre and 1200 ft^3/acre would compete for the same node in the network. The key in using state neighborhoods is making sure that your intervals are not so large as to violate the principle of optimality (Pelkki 1997).

AN EXAMPLE OF DP AND FIELD TESTING - YELLOW POPLAR MANAGEMENT

Description of NESTER

Nester (NEighborhood STate EvaluatoR) is a forward recursive dynamic programming computer program for Windows-based personal computers (Pelkki 1997). It utilizes the GROW subroutine (Brand 1981) from the TWIGS individual-tree growth projection system (Miner and others 1988). NESTER allows the user to choose any combination of six state variables (cubic foot volume per acre, number of trees per acre, average diameter, basal area per acre, number of thinnings, and average tree grade per acre). Stage intervals can be as short as one year or as long as 30 years. As NESTER utilizes an individual-tree growth model, stumpage prices must be provided for each species group included in the stand. NESTER can be used to model mixed species stands, however, at the present time, there are no thinning algorithms that select by species. NESTER does project tree grade using the method presented by Yaussy (1993). Also, because of the individual-tree growth model, NESTER can simulate many types of thinning operations, including mechanical thinning, thinning from above, thinning from below, and thinnings based on tree quality (improvement thinning and high-grading). NESTER can simulate constant, real price increases, or stochastic price changes independently for sawtimber and other roundwood size classes.

NESTER Studies on Yellow Poplar

Two studies have been published to date using NESTER to study yellow-poplar (Pelkki and Arthaud 1998, Pelkki 1999). One study focused upon changing markets and prices and their effects on yellow-poplar management, the other examined changing thinning strategies and the impact on financial returns from yellow- poplar. The results of these studies indicate the strength of dynamic programming for sensitivity analysis and the overall findings of these papers will be compared to published results of field studies reported in Beck and Della-Bianca (1981). The complete results of those studies are not included here. This paper merely highlights some of the new information obtained through computer simulations and identifies some areas where additional field-based work is needed.

Input Data for NESTER

Initial stand conditions were derived from published diameter distributions (McGee and Della-Bianca 1967, Knoebel and others 1986) for 20-year-old, average-stocked, yellow-poplar stands with site indices of 90, 110,

and 130 ft at 50 years. Potential grade distributions (Hanks 1976) representing low, average, and high stem quality classes were defined using U.S. Forest Service Forest Inventory and Analysis data for yellow-poplar in the region. Thus, nine different combinations of site index and stem quality distribution served as starting points for the DP simulations.

Stumpage prices were obtained from regional price reports and from Timber Mart South (1994). Capital costs were generally set at 4 percent. Sensitivity to real price increases and higher and lower costs of capital were investigated.

NESTER runs were completed using two states, number of trees per acre and cubic foot volume per acre, with the state neighborhoods of 10 trees per acre and 10 ft^3 per acre, respectively. The stage interval was set at 2 years, but additional runs were investigated using different stage intervals. All thinning options but mechanical thinning were simulated (thinning from above, thinning from below, thinning from above and below, high-grading, and improvement thinning) at intensities ranging from 10 percent to 50 percent of the initial basal area in 5 percent increments. Thus, from each state, 47 options were simulated. The state network represented 46^n+1 states (where n = stage). The initial stage was age 20, and each stage increased by two year increments. Therefore, stage 11 (age = 42) represented $46^{11}+1$ possible ways to manage a stand from age 20 to age 42.

Comparing Results of NESTER-Based Studies to Previous Work

When thinning yellow-poplar stands, Beck and Della Bianca (1981) note that cultural work in sapling and pole-sized stands is very costly and there are few markets for this material. When thinnings are done, Beck and Della-Bianca (1981) recommend thinnings from below to concentrate the value on larger, high-value stems. They also noted that intermediate yellow-poplar trees would respond to thinning from above. In a recently completed DP study, Pelkki (1999) found that improvement thinnings, while earning negative to very low initial returns (-$5.1 to $31.5 per acre) were commonly part of the optimal financial stand regime. These improvement thinnings removed first large trees of low quality, then very small trees, and finally low to higher grade factory sawtimber grade trees until a basal area target was reached. In this manner, the trees with the most valuable future growth potential were retained.

In stands with high-quality stem distributions, the DP-based studies (Pelkki and Arthaud 1998, Pelkki 1999) included thinnings that removed 40-50 percent of the basal area. Furthermore, some of the optimal financial regimes included 4-6 thinnings (not all at the 40-50 percent intensity level). Beck and Della-Bianca (1981) report considerable leeway in manipulating yellow-poplar stocking levels to achieve diameter growth and quality goals without sacrificing volume production. However, the frequency and intensity of the DP-based harvest may initiate advance regeneration which would increase the final harvest and site regeneration costs.

In studies reported in Beck and Della-Bianca (1981), first thinnings are recommended as early as 15-20 years and can be repeated every 5-15 years over the rotation. The DP-based studies (Pelkki and Arthaud 1998, Pelkki 1999) have an initial stand state that is 20 years of age and so cannot simulate thinnings before the age of 22 years (1st stage). However, in the DP-studies, most initial thinnings occurred between the ages of 22 and 28 years. Multiple thinnings in the DP-based studies ranged from 2-12 years apart with most being 4-10 years apart.

In regimes with multiple thinnings, Beck and Della-Bianca (1981) recommend that later thinnings be lighter because basal area growth response in older stands is lower. The DP-studies (Pelkki and Arthaud 1998, Pelkki 1999) had heavier later thinnings, possibly due to the financial, rather than volume objectives of the simulation. Both Beck and Della-Bianca (1981) and the DP-based studies (Pelkki and Arthaud 1998, Pelkki 1999) found that intense thinnings shortened the rotation.

While not explicitly discussed in Beck and Della-Bianca (1981), the yield tables suggest an optimal rotation of 50-70 years for fiber production. The DP-based studies (Pelkki and Arthaud 1998, Pelkki 1999) found maximum financial returns with rotations as short as 32 years or as long as 66 years. Factors contributing to shorter rotations were a good market for pulpwood, OSB, and other fiber-based products, high interest rates, and higher site indices. Factors contributing to longer rotation lengths were high stem quality distributions, lower interest rates, and higher sawtimber prices (or an absence of sub-sawtimber markets). These factors favored longer, sawtimber-focused rotations.

For yellow-poplar management, the DP-based studies uncover three major issues that need additional research in a field-based setting. First and foremost, can improvement thinnings, based on potential tree grading bring the economic returns projected by the computer simulations? Additional studies have shown (Pelkki and Ringe 1998) that the earlier a valid potential tree grade can be applied, the greater the economic returns. Secondly, are harvest costs in poletimber stands prohibitive? The DP-simulations assume a fixed entry cost on all thinnings and a 15 percent thinning penalty on all stumpage prices (to reflect the added cost of thinning over a final harvest). They also charge harvest costs for pre-merchantable stems at the rate of 10 percent of the merchantable price (if the stems were of merchantable size). Studies by Kluender and others (1996) suggest that harvest cost of small diameter stems are greater than proportional to volume in southern pine stands. With the increase in markets for hardwood pulpwood throughout the central Appalachian region, thinning and site preparation cost studies would appear necessary. Finally, the DP-based simulations suggest that multiple thinning regimes with 4-6 entries prior to final harvest may optimize financial returns in some yellow-poplar stands. Some field-based tests of such regimes

would answer questions related to problems in stand integrity and advance regeneration treatment costs.

SUMMARY - WHY BOTH APPROACHES ARE NEEDED

Unquestionably, both field-based and simulation-based approaches are necessary for forest researchers in the future. Field-based studies provide on-the-ground confirmation of new practices and working methods for field application. They cannot, however, respond quickly to new markets or changing economic conditions. While computer simulations can perform thorough sensitivity analyses to many variables, they lack the real world operational constraints that are often too complex to model and may even be unanticipated by the researcher. Computer simulations are best left to exploratory research with their results confirmed by field-based studies which can then lead to changes in actual management practices.

REFERENCES

Arthaud, G.J. 1986. Economic comparisons of thinning from above and below in loblolly pine plantations using dynamic programming. Blacksburg, VA: Virginia Polytechnic Institute and State University. 139 p. M.S. thesis.

Beck, D.E.; Della-Bianca, L. 1970. Yield of unthinned yellow-poplar. Res. Pap. SE-58. U.S. Department of Agriculture, Forest Service. 20 p.

Beck, D.E.; Della-Bianca, L. 1975. Board-foot growth and yield of yellow-poplar after thinning. Res. Pap. SE-123. U.S. Department of Agriculture, Forest Service. 20 p.

Beck, D.E.; Della-Bianca, L. 1981. Yellow-poplar: characteristics and management. Agric. Handb. 583. U.S. Department of Agriculture, Forest Service. 91 p.

Brand, G.J. 1981. "GROW" - A computer subroutine that projects the growth of trees in Lake States' forests. Res. Pap. NC-207. U.S. Department of Agriculture, Forest Service.

Brodie, J.D.; Kao, C. 1979. Optimizing thinning in Douglas-fir with three descriptor dynamic programming to account for accelerated diameter growth. Forest Science. 25: 665-672.

Dykstra, D.P. 1984. Mathematical programming for natural resource management. McGraw-Hill. 318 p.

Hanks, L.F. 1976. Hardwood tree grades for factory lumber. Res. Pap. NE-333. U.S. Department of Agriculture, Forest Service. 81 p.

Kluender, R.; Lortz, D.; McCoy, W. [and others]. 1996. Harvesting time, costs, and productivity in Arkansas. In: Caulfield, J.; Bullard, S., eds. Proceedings of 1995 southern forestry economics workshop; April 17-19; New Orleans, LA: 260-274.

Knoebel, B.R.; Burkhart, H.E.; Beck, D.E. 1986. A growth and yield model for thinned stands of yellow-poplar. For. Sci. Monograph 27. 62 p.

McGee, C.E.; Della-Bianca, L. 1967. Diameter Distributions in natural yellow-poplar stands. Res. Pap. U.S. Department of Agriculture, Forest Service. 7 p.

Miner, C.L.; Walters, N.R.; Belli, M.L. 1988. A guide to the TWIGS program for the North Central United States. Gen. Tech. Rep. NC-125. U.S. Department of Agriculture, Forest Service.

Pelkki, M.H. [In press]. Impacts of thinning strategies on financial returns from yellow-poplar stands. Southern Journal of Applied Forestry.

Pelkki, M.H. 1997. The effects of neighborhood storage size on dynamic programming solutions. Forest Science. 43(3): 387-395.

Pelkki, M.H.; Arthaud, G.J. 1998. Evaluating changes in yellow-poplar management in response to new low-end markets. Northern Journal of Applied Forestry. 15(1): 37-42.

Rauscher, H.M.; Young, M.J.; Webb, C.D. 1997. Testing the accuracy of growth and yield models for southern hardwood forests. Unpublished manuscript. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station, Bent Creek Research and Demonstration Forest.

Timber Mart South. 1994. First and second quarter reports. Athens, GA: University of Georgia. Vol. 19.

Yaussy, D. 1993. NE-TWIGS update: incorporating tree quality. The Compiler. Vol. 11(1): 8-12.

USE OF GPS AND GIS IN HARDWOOD FOREST INVENTORY

C.J. Liu[1]

Abstract—Recent advances in GPS satellite survey, geographic information systems, and a portable electronic distance measurement device are useful new tools which, when combined with classical tree measurement, timber volume calculation, statistical inventory procedures, provide a pathway for generating highly accurate digital timber stand information for effective forest management decision makings. This presentation describes the methods and procedures used in the design of an intensive timber inventory conducted in a mountainous forest watershed.

In this prototype study, GPS, laser ranging, and GIS techniques are used to replace traditional forest inventory tools such as chains, compasses, altimeters, and optical forks. GPS land navigation technique is employed to locate inventory plot centers and a differential GPS point positioning technique is used to permanently determine surveying plot locations. A laser ranging instrument, called Criterion, is used to measurement upper-stem diameters and subsequently to determine merchantable lengths of commercial trees. Georeferenced timber survey data and computer generated environment data such as slope and aspect maps are stored in a geographic information system for subsequent analyses.

The inventory uses a traditional systematic line-plot design to tally timbers in an eastern hardwood forest situated on the western slope of the Appalachian mountain. Using Criterion's laser ranging capability, plot trees are identified by measuring the distance between plot and tree centers. For a plot tree, its dbhob is measured by a caliper and its merchantable length is measured by the criterion. The Criterion's diameter and height measurement functions allow foresters to accurately determine the position of a prescribed upper-stem diameter and the measure the vertical distance between two points on tree stems. The digital display feature provided by this instrument eliminates the need to interpret analogous scales on conventional tree measurement and portable surveying equipment. Tree measurement data are subsequently inputted into a personal computer for the computations of both board-foot and cubic-foot saw-timber volumes. All these measurements exceed functional accuracies specified for intended forest inventory work.

Using statistic estimation methods, plot data are expanded to provide interval estimates of timber stocking on a per acre basis or for the entire forest stand. In addition, all inventory data are entered into a GIS framework which contains environmental data such as aspect and slope maps. The GIS is used to store, display, and query of timber stands information for effective forest management decision makings.

---

[1] Department of Forestry, University of Kentucky, Lexington, KY 40546.

Citation for proceedings: Stringer, Jeffrey W.; Loftis, David L., eds. 1999. Proceedings, 12th central hardwood forest conference; 1999 February 28-March 1-2; Lexington, KY. Gen. Tech. Rep. SRS-24. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 293 p. [Oral presentation abstract].

USE JAVA AND THE INTERNET TO MANAGE DATA AND PREDICT THE FUTURE OF FOREST STANDS

J.J. Colbert and George Racin[1]

Abstract—A Java based software package provides a user with Internet access the ability to store, summarize, and translate field collected data describing forest stands and to run simulations to assess possible future scenarios that include stand management and the effects of gypsy moth defoliation. The model interface is constructed as a Java applet that will run within a Java-compliant Internet world-wide-web browser, allowing a user from any computer that has Internet and web access to use these tools. The model user also has the ability to access data from a large collection of example stands that are available from a server-side database. If the user has a locally available Java Virtual Machine (JVM), then it is possible to use a Java application version of these software programs that permits the local storage, access, and management of input and output data files generated through use of the program. Data from field plots entered through the user interface are summarized on a per acre basis and from these data, the user can directly obtain estimates of stocking, volume, and value, as well as stand structure and related habitat data summaries. Simulations permit the user to look at potential growth scenarios under hypothesized management actions.

[1] U.S. Department of Agriculture, Forest Service, Morgantown, WV 26505-3101.

Citation for proceedings: Stringer, Jeffrey W.; Loftis, David L., eds. 1999. Proceedings, 12th central hardwood forest conference; 1999 February 28-March 1-2; Lexington, KY. Gen. Tech. Rep. SRS-24, Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 293 p. [Oral presentation abstract].

ESTIMATING PREVIOUS DIAMETER FOR INGROWTH TREES ON REMEASURED HORIZONTAL POINT SAMPLES

Susan L. King and Stanford L. Arner[1]

Abstract—The purpose of this study was to develop an improved procedure to estimate a previous diameter for ongrowth and nongrowth trees on plots sampled by variable radius plot sampling. However, the models built can be used whenever a previous diameter is required and the appropriate independent variables are available. Data for this study were from 1,965 remeasured forest inventory plots in the 1989 inventory of West Virginia conducted by the Northeastern Research Station's Forest Inventory and Analysis (NEFIA) unit. The investigation focused on three areas. First, we investigated whether breaking the data into six groups based on the rank average diameter growth was superior to NEFIA's procedure of breaking the data into seventeen species groups. Second, we investigated whether basal-area increment (BAI) was superior to diameter increment (DI) as a dependent variable. Finally, we investigated additional independent variables. Based on the R^2, mean residual, mean absolute error, and root mean square error, the best model had six species groups, BAI as the dependent variable, and a slightly expanded set of independent variables.

INTRODUCTION

Missing observations are common in "real world" data sets. In forestry, there are many instances in which a diameter for a previous period is required but not available. The impetus for this study was a problem encountered by the Forest Service's Forest Inventory and Analysis (FIA) units which are required to produce periodic tables showing net volume change from the previous inventory as well as the components of change: ingrowth (I), survivor growth (S), mortality (M), and cut (C). Equation (1) expresses net change as the sum of these components:

$$V_2 - V_1 = I + S - M - C \qquad (1)$$

where:

V_1 = volume at time period 1;

V_2 = volume at time period 2.

Martin (1982) defined six possible categories of trees encountered on remeasured point samples. The first four, ingrowth, survivor, mortality, and cut trees were measured at the first inventory. The remaining categories, ongrowth and nongrowth trees, were alive and included only in the second inventory. Ongrowth trees were nonmerchantable (below minimum dbh) at the first inventory but of merchantable size at the second inventory. Nongrowth trees were above minimum dbh at the first inventory, but grew sufficiently to be included in the second inventory.

Most FIA units have used multiple subplot horizontal point samples (prism plots) to inventory trees larger than a specified merchantable diameter. With this type of sample, the probability of selecting a sample tree is proportional to the basal area of the individual tree and depends on the basal-area factor of the prism.

Traditional estimation procedures for total net change and its components yielded estimates for which the two sides of equation (1) do not agree because ongrowth and nongrowth trees were excluded as components of growth. Martin (1982) obtained compatible estimators by including in ingrowth both ongrowth and nongrowth trees. Van Deusen and others (1986) improved the compatible estimators by rearranging Martin's equation to include nongrowth with survival growth. They showed that the standard error of this new estimator of survival growth was smaller than Martin's traditional estimator. Also, the estimator for ingrowth used by Van Deusen and others cannot be negative. Roesch and others (1989) showed that the estimation of survival growth could be improved with additional rearrangement of some components. To distinguish between ongrowth and nongrowth trees, these new procedures require the estimation of previous diameter for trees that were not measured at the initial inventory.

NEFIA unit has used several plot designs, including a 10-point variable radius design. To obtain compatible estimates of change with these plots, we need to estimate previous diameter for trees measured only at time period 2.

To estimate previous diameter, NEFIA currently uses trees measured at both inventories to develop regression equations. The model is:

$$DI = f(DBH2, TRCLS2, CRNCLS2, CRATIO2, CRCC2, DCR2) \qquad (2)$$

where:

DI = annual diameter increment;

DBH2 = tree diameter at time period 2;

TRCLS2 = tree class at time period 2, a measure of tree quality;

[1] Operations Research Analyst and Supervisory Forester, respectively, USDA Forest Service, 100 Matsonford Road, 5 Radnor Corp Ctr Ste 200, Radnor, PA 19087-4585.

Citation for proceedings: Stringer, Jeffrey W.; Loftis, David L., eds. 1999. Proceedings, 12th central hardwood forest conference; 1999 February 28-March 1-2; Lexington, KY. Gen. Tech. Rep. SRS-24. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 293 p. [Peer-reviewed paper].

CRNCLS2 = crown class at time 2, a measure of crown position in the canopy;

CRATIO2 = crown ratio at time 2, the proportion of a tree with a live crown;

CRCC2 = CRATIO2/CRNCLS2;

DCR2 = DBH2 · CRATIO2.

The data are divided into seventeen species groups. For each group, a stepwise backward elimination regression procedure finds the functional relationship between the dependent and independent variables. Annual DI is the dependent variable as opposed to total increment between inventories because the trees are remeasured at different intervals. The procedures developed by Roesch and others (1989) and Van Deusen and others (1986) assume that good estimates of previous diameter for ongrowth and nongrowth trees are possible. With the NEFIA procedure, the R^2 for each of the seventeen species groups is low, *e.g.*, the most recent inventory from West Virginia had R^2 values ranging from 0.06 to 0.56.

PROCEDURE

The data for this study are from 1,965 remeasured forest inventory plots in West Virginia (DiGiovanni 1990). Trees were initially measured in 1975 and remeasured in 1988. Only trees larger than 5 inches in diameter at both inventories were used. The data were split into a model data set with 8,723 observations and a validation data set with 7,951 observations.

Using multiple linear regression, we investigated different sets of independent variables, different species-group compositions, and BAI as a dependent variable. Although the explanatory variables have biological meaning with respect to competitive position, size, and tree quality, the model itself is not easily interpreted biologically as is the case with models developed by Teck and Hilt (1991) and Quicke and others (1994). However, our objective was to find a better estimate of annual DI for trees alive at both inventories. It was not necessary to account for ingrowth, cut, or mortality.

Apart from the variables used by NEFIA, we investigated additional variables reported in the literature as important in modeling tree growth. The additional independent variables were restricted to those measured by NEFIA and those that could be calculated:

BA2 = total plot basal area of trees at least 5 inches in diameter at time 2;

TPA2 = total number of trees per acre of trees at least 5 inches in diameter at time 2;

MD2 = plot medial diameter of trees at least 5 inches in diameter at time 2;

QD2 = plot quadratic diameter of trees at least 5 inches in diameter at time 2;

BAL2 = total basal area of trees larger than the sample tree at time 2;

RLD2 = relative diameter; ratio of tree diameter to medial diameter of plot, MD2;

RLQD2 = relative diameter; ratio of tree diameter to quadratic diameter of plot, QD2.

The plot variables, BA2, TPA2, MD2, and QD2, are expressions of site occupancy or total competition and size of the trees on the plot. The tree variables, BAL2, RLD2, and RLQD2, are measures of the competitive position of the sample tree relative to other trees on the plot.

The basal area larger (BAL2) variable is used in the potential growth times modifier type of model found in the NC and NE-TWIGS growth simulators (Hilt and Teck 1989: Miner and others 1988: Quicke and others 1994). Marquis (1991) used relative diameter to model diameter growth.

Variables not considered are tree or stand age and a measure of site productivity. NEFIA does not determine the age of individual trees, and stand age was excluded because many of the plots are classified as uneven-aged. Also, experience has shown a poor relationship between diameter growth and the site-productivity measure determined by NEFIA.

We investigated both DI and BAI as the dependent variable. Because of unequal number of years between plot measurements, annual increment was modeled as:

$$DI = \frac{(DBH2 - DBH1)}{N} \qquad (3)$$

or

$$BAI = \frac{(BA2 - BA1)}{N} \qquad (4)$$

where:

N = number of years between measurements on the plot;

DBH1 = tree diameter at time period 1;

DBH2 = tree diameter at time period 2;

BA1 = tree basal area at time period 1;

BA2 = tree basal area at time period 2.

The seventeen species groups used by NEFIA are based on form class and are used in our volume equations (Scott 1979). To investigate whether another grouping of species is more appropriate to model diameter growth, we formed six species groups based on rank of the average diameter growth for a species. Mean annual diameter growth ranged from 0.068 for the lowest ranked group to 0.203 for the highest.

RESULTS

We compared the R^2 values, mean residuals, mean absolute residuals, and root mean square errors using the NEFIA independent variables for the total sample of both the model and validation data sets (Table 1). The comparison is for six procedures: 1) the stepwise procedure on each of seventeen species groups using DI as the dependent variable; 2) all variables for each group using DI; 3) all variables ignoring species group using DI;

Table 1—R^2s, mean residuals, mean absolute residuals, and root mean square errors (MSE) for the stepwise procedure using NEFIA variables with three species groupings

Procedure	R^2	Mean residual	Mean abs. residual	Root MSE
		Model data set		
Diameter increment				
Stepwise, 17 groups	0.348	0.0	0.0530	0.0692
Nonstepwise, 17 groups	0.349	0.0	0.0529	0.0691
Nonstepwise, no groups	0.211	0.0	0.0584	0.0762
Nonstepwise, 6 groups	0.351	0.0	0.0530	0.0691
Basal-area increment				
Nonstepwise, 17 groups	0.557	0.0033	0.0440	0.0571
Nonstepwise, 6 groups	0.565	0.0013	0.0436	0.0566
		Validation data set		
Diameter increment				
Stepwise, 17 groups	0.312	0.0017	0.0544	0.0721
Nonstepwise, 17 groups	0.312	0.0017	0.0545	0.0721
Nonstepwise, no groups	0.207	0.0022	0.0585	0.0775
Nonstepwise, 6 groups	0.322	0.0005	0.0545	0.0716
Basal-area increment				
Nonstepwise, 17 groups	0.536	0.0032	0.0451	0.0592
Nonstepwise, 6 groups	0.553	0.0017	0.0448	0.0582

4) all variables using six species groups and DI; 5) all variables using seventeen species groups with BAI as the dependent variable; and 6) all variables using six species groups and BAI. Equation (5) and (6) give the formulas for calculating the R^2 and root mean square errors, respectively, for the combined data.

$$R^2 = 1 - \left(\frac{SSRES}{SSTOT}\right) \quad (5)$$

$$RMSE = \sqrt{\frac{SSRES}{n_i}} \quad (6)$$

where:

$$SSRES = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(Y_{ij} - \hat{Y}_{ij}\right)^2;$$

$$SSTOT = \sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(Y_{ij} - \overline{Y}\right)^2;$$

Y_{ij} = observed DI;

$\hat{Y}_{ij}$ = predicted DI;

$\overline{Y}$ = overall mean DI;

n_i = number of trees in group i;

g = number of species in a group, 17, 6, or 1.

The predicted value for the validation data set, $\hat{Y}_{ij}$, was obtained using the coefficients estimated with the model data set.

The stepwise procedure for the seventeen species groups offers little advantage over the nonstepwise procedure for each group. There is little difference in any of the statistics between these two procedures. Each of the NEFIA independent variables was significant in several of the groups, indicating that eliminating one or more of the variables would be inappropriate. Yet there is an advantage in grouping the species. All of the statistics for both the model and validation data sets show improvement when the data are divided into species groups. However, there is little difference between the results using the seventeen groups based on form class and the six groups based on rank of mean DI. As seen in Table 1, there is a marked improvement when BAI replaces DI as the dependent variable. For the BAI model, all results are expressed in reference to DI using the translation:

$$\hat{DI}(BA) = \frac{\sqrt{\frac{(BA1 + N \cdot \hat{BAI})}{K}} - DBH1}{N} \quad (7)$$

where:

$\hat{BAI}$ = predicted BAI;

K = 0.005454154, a conversion factor from diameter in inches to basal area in square feet.

All of the statistics use $\hat{DI}(BA)$ as the predicted DI for the basal-area model.

Table 2 presents a more complete analysis of the differences between results using DI versus BAI for the

Table 2—Comparison statistics for the diameter-increment and basal-area increment models for model and validation data sets

Species group	No. of trees	Diameter-increment model: Mean DI	R^2	Mean residual	Mean abs. residual	Root MSE	Basal-area increment model: R^2	Mean residual	Mean abs. residual	Root MSE
				Model data set						
1	247	0.0681	0.045	0.0	0.0322	0.0408	0.235	0.0011	0.0288	0.0365
2	947	0.0929	0.085	0.0	0.0415	0.0525	0.315	0.0004	0.0359	0.0454
3	2489	0.1145	0.157	0.0	0.0434	0.0562	0.370	0.0012	0.0374	0.0486
4	2371	0.1434	0.201	0.0	0.0556	0.0714	0.471	0.0010	0.0452	0.0582
5	783	0.1595	0.219	0.0	0.0564	0.0720	0.473	0.0021	0.0464	0.0591
6	1876	0.2026	0.212	0.0	0.0689	0.0880	0.513	0.0022	0.0540	0.0692
All groups	8723	0.1416	0.351	0.0	0.0528	0.0691	0.565	0.0013	0.0435	0.0566
				Validation data set						
1	97	0.0700	0.001	0.0018	0.0337	0.0413	0.204	0.0022	0.0300	0.0369
2	616	0.0953	0.079	0.0024	0.0425	0.0546	0.315	0.0029	0.0368	0.0471
3	2417	0.1146	0.183	0.0001	0.0436	0.0557	0.392	0.0017	0.0373	0.0481
4	2448	0.1435	0.161	0.0016	0.0561	0.0732	0.450	0.0018	0.0455	0.0592
5	509	0.1662	0.115	0.0043	0.0598	0.0788	0.401	0.0027	0.0497	0.0649
6	1864	0.2030	0.201	-0.0021	0.0700	0.0898	0.514	0.0014	0.0551	0.0701
All groups	7951	0.1455	0.322	0.0005	0.0545	0.0716	0.552	0.0018	0.0447	0.0582

procedure with six-species groups. Although the bias (as expressed by the mean residual) is slightly greater for the BAI model, the R^2, mean of absolute residuals, and root mean square error are substantially smaller. Mean annual diameter growth (mean DI) for each species group also is included in Table 2.

Table 3 gives the results for different combinations of independent variables for the six species groups. Some of the independent variables discussed previously were added to the NEFIA variables. The models listed are the best 1, 2, and 3 variable combinations. The additional variables express plot occupancy and tree competitive position. In the remaining three models, CRCC2 and CRNCLS2 were deleted. In each of the three models, BAL2, RLD2, or RLQD2 was substituted for CRNCLS2. These variables measure a tree's competitive position relative to the other trees on a plot.

The results for the models with variables substituted for or added to the FIA variables are mixed. Each of the four statistics presented in Table 3 indicate a different "best" model. All differences are small and no model represents a substantial improvement over the others.

Table 3—Comparison statistics for models with variables substituted for or added to NEFIA variables using six-group BAI model and model data sets

Variable	R^2	Mean residual	Mean abs. residual	Root MSE
NEFIA[a]	0.565	0.0013	0.0436	0.0566
DBH2, TRCLS2, CRATIO2, DCR2, BAL2	0.575	0.0014	0.0431	0.0559
DBH2, TRCLS2, CRATIO2, DCR2, RLD2	0.562	0.0010	0.0438	0.0567
DBH2, TRCLS2, CRATIO2, DCR2, RLQD2	0.567	0.0010	0.0435	0.0564
NEFIA + BAL2	0.578	0.0016	0.0429	0.0557
NEFIA + BAL2 + BA2	0.581	0.0015	0.0428	0.0555
NEFIA + BAL2 + BA2+TPA2	0.581	0.0015	0.0428	0.0555

[a] NEFIA= DBH2, TRCLS2, CRNCLS2, CRATIO2, CRCC2, DCR2.

We used the BAI model with six species groups and NEFIA variables to determine how the model performed over the range of tree sizes. The mean residual, mean absolute residual, and root mean square error were determined for 2-inch diameter classes (Table 4). As indicated by the mean residual, for the validation data set there is a small average over prediction for diameters less than 9 inches and greater than or equal to 23 inches. For the diameter classes in between there is a small underprediction.

DISCUSSION AND CONCLUSIONS

With our method there was a marked improvement over current NEFIA procedures in all model valuation statistics except for the mean residual. For the validation data set, the change in the dependent variable from annual DI to annual BAI produced a 72 percent increase in R^2, a 17 percent decrease in mean absolute residual, and an 18 percent decrease in root mean square error.

Grouping by species was better than not grouping them. The grouping based on ranked diameter growth showed only a slight improvement over the seventeen NEFIA form class groups for both the DI and BAI models. Other groupings were not investigated.

There was only a slight improvement when even the best of the other procedures and models was used. As a result, there seems little reason to find a "best" set of independent variables for each group using a stepwise procedure as the outcome using all variables for all groups is essentially the same.

The variables considered as substitutions for or additions to the NEFIA variables produced little improvement. Approximately the same results could be obtained using any of the different combinations of NEFIA and the other plot and tree variables considered here.

Quicke and others (1994) developed a biologically interpretable model using a function with potential growth multiplied by a growth modifier for a single species with plots chosen for specified characteristics. For this study, trees are located on plots chosen at random from a wide range of forest conditions. Disturbance on the plot was not taken into account. There were 78 species and a wide range in size. The results presented here closely approximate those obtained by Quicke and others (1994).

Data used by Teck and Hilt (1991) are from the same type of unstructured design as those presented in this study. They also developed and used a biologically interpretable DI model. With their validation data set, the overall mean prediction error was 0.013 and the root mean square error was 0.085. These statistics are substantially higher than ours.

The coefficients of the models developed here can be used for purposes other than that of NEFIA if the independent variables included in the model are measured. Table 5 contains coefficients for the six-species-group BAI model using NEFIA variables plus BAL2. Table 6 lists the tree species assigned to each of six species groups.

The coefficients developed from this study should be applied to other regions with caution.

Table 4—Comparison statistics by diameter class for six-group basal-area increment model

	Model data set					Validation data set				
Dbh class	No. of trees	Mean DI	Mean Residual	Mean abs. residual	Root MSE	No. of trees	Mean DI	Mean residual	Mean abs. residual	Root MSE
In.										
5 - 6.9	529	0.0573	-0.0043	0.0231	0.0293	490	0.0565	-0.0034	0.0230	0.0289
7 - 8.9	1353	0.0930	-0.0018	0.0313	0.0394	1136	0.0965	-0.0010	0.0323	0.0398
9 - 10.9	1371	0.1246	0.0025	0.0412	0.0518	1202	0.1259	0.0028	0.0401	0.0508
11 - 12.9	1316	0.1401	0.0011	0.0458	0.0577	1179	0.1466	0.0042	0.0454	0.0592
13 - 14.9	1250	0.1587	0.0049	0.0468	0.0614	1087	0.1576	0.0026	0.0487	0.0636
15 - 16.9	946	0.1676	0.0008	0.0483	0.0615	918	0.1741	0.0046	0.0506	0.0651
17 - 18.9	702	0.1787	0.0027	0.0532	0.0680	643	0.1901	0.0089	0.0543	0.0696
19 - 20.9	470	0.1903	0.0064	0.0532	0.0686	480	0.1845	0.0006	0.0564	0.0719
21 - 22.9	296	0.1909	0.0007	0.0554	0.0711	263	0.1882	0.0003	0.0525	0.0665
23 - 24.9	180	0.2018	0.0056	0.0532	0.0698	222	0.1921	-0.0044	0.0580	0.0709
> 25	310	0.1913	-0.0044	0.0499	0.0614	331	0.1851	-0.0103	0.0532	0.0641
All trees	8723	0.1416	0.0013	0.0436	0.0566	7951	0.1455	0.0017	0.0448	0.0582

Table 5—Regression coefficients for six-group basal-area increment model using NEFIA variables and BAL2

Species group	Intercept	DBH2	TRCLS2	CRATIO2	CRNCLS2	DCR2	CRCC2	BAL2
1	0.00107	0.00065	0.00011	0.00028	-0.00043	0.00005	-0.00059	-0.000014
2	0.01448	0.00028	-0.00127	0.00053	-0.00225	0.00020	-0.00367	-0.000033
3	0.02058	0.00064	-0.00309	-0.00115	-0.00272	0.00025	-0.00301	-0.000042
4	0.02239	0.00048	-0.00104	-0.00159	-0.00365	0.00041	-0.00486	-0.000048
5	-0.00150	0.00160	-0.00132	-0.00306	0.00041	0.00021	0.00582	-0.000055
6	0.00132	0.00214	-0.00405	-0.00458	0.00219	0.00016	0.00957	-0.000110

Table 6—Tree species assigned to one of six species groups based on rank of mean diameter increment

Species group	Common name	Scientific name
1	Hawthorn	*Crataegus sp.*
	Shortleaf pine	*Pinus echinata*
	Flowering dogwood	*Cornus flordia*
	European alder	*Alnus glutinosa*
	Table mountain pine	*Pinus pungens*
	Black willow	*Salix nigra*
	Chinkapin oak	*Quercus muehlenbergii*
	Eastern hophornbeam	*Ostrya virginiana*
	Silver maple	*Acer saccharinum*
	Blackgum	*Nyssa sylvatica*
	Willow oak	*Quercus phellos*
	Sourwood	*Oxydendrum arboreum*
2	Maple sp.	*Acer sp.*
	Quaking aspen	*Populus tremuloides*
	Pitch pine	*Pinus rigida*
	Overcup oak	*Quercus lyrata*
	Post oak	*Quercus stellata*
	Red spruce	*Picea rubens*
	Pin cherry	*Prunus pennsylvanica*
	Am. hornbeam, Musclewood	*Carpinus caroliniana*
	Osage-orange	*Maclura pomifera*
	River birch	*Betula nigra*
	Yellow birch	*Betula alleghaniensis*
	Basswood sp.	*Tilia sp.*
	Black walnut	*Juglans nigra*
	Virginia pine	*Pinus virginiana*
	Mockernut hickory	*Carya tomentosa*
	Sweet birch	*Betula lenta*
	Common persimmon	*Diospyros virginiana*
3	Bur oak	*Quercus macrocarpa*
	Ohio buckeye	*Aesculus glabra*
	Pignut hickory	*Carya glabra*
	Hickory sp.	*Carya sp.*
	American beech	*Fagus grandifolia*
	Magnolia sp.	*Magnolia sp.*
	Chokecherry	*Prunus virginiana*
	Buckeye, Horsechestnut	*Aesculus sp.*
	Chestnut oak	*Quercus prinus*
	Sassafras	*Sassafras albidum*
	White basswood	*Tilia heterophylla*
	Shagbark hickory	*Carya ovata*
	Black locust	*Robinia pseudoacacia*
4	Hackberry	*Celtis occidentalis*
	White oak	*Quercus alba*
	Apple sp.	*Malus sp.*
	Eastern hemlock	*Tsuga canadensis*
	Bitternut hickory	*Carya cordiformis*
	Yellow buckeye	*Aesculus octandra*
	Slippery elm	*Ulmus rubra*
	Swamp white oak	*Quercus bicolor*
	Cucumbertree	*Magnolia acuminata*
	American basswood	*Tilia americana*
	Red maple	*Acer rubrum*
	American elm	*Ulmus americana*
	Sugar maple	*Acer saccharum*
	Shellbark hickory	*Carya laciniosa*
	Striped maple	*Acer pennsylvanicum*
	Eastern red-cedar	*Juniperus virginiana*
5	Black oak	*Quercus velutina*
	White ash	*Fraxinus americana*
	Butternut	*Juglans cinerea*
	Elm sp.	*Ulmus sp.*
	Sycamore	*Platanus occidentallis*
	Green ash	*Franxinus pennsylvanica*
	Boxelder	*Acer negundo*
	Black cherry	*Prunus serotina*
	Eastern white pine	*Pinus strobus*
	Scarlet oak	*Quercus coccinea*
6	Northern red oak	*Quercus rubra*
	Bigtooth aspen	*Populus grandidentata*
	Yellow-poplar	*Liriodendron tulipifera*
	Pin oak	*Quercus palustris*
	Prunus sp.	*Prunus sp.*
	Ailanthus	*Ailanthus altissima*
	Black maple	*Acer nigrum*
	Eastern redbud	*Cercis canadensis*
	Southern red oak	*Quercus falcata v. falcata*

REFERENCES

DiGiovanni, D.M. 1990. Forest statistics for West Virginia—1975 and 1989. Resour. Bull. NE-114. U.S. Department of Agriculture, Forest Service. 172 p.

Hilt, D.E.; Teck, R.M. 1989. NE-TWIGS: An individual-tree growth and yield projection system for the Northeastern United States. The Compiler. 7(2): 10-16.

Marquis, D.A. 1991. Independent effects and interactions of stand diameter, tree diameter, crown class, and age on tree growth in mixed-species, even-aged hardwood stands. In: Proceedings, 8th central hardwood forest conference. Gen. Tech. Rep. NE-148. U.S. Department of Agriculture, Forest Service: 442-458

Martin, G.L. 1982. A method for estimating ingrowth on permanent horizontal sample points. Forest Science. 28(1): 110-114.

Miner, C.L.; Walters, N.R.; Belli, M.L. 1988. A guide to the TWIGS program for the North Central United States. Gen. Tech. Rep. NC-125. U.S. Department of Agriculture, Forest Service. 105 p.

Quicke, H.E.; Meldahl, R.S.; Kuch, J.S. 1994. Basal area growth of individual trees: a model derived from a regional longleaf pine growth study. Forest Science. 40(3): 528-542.

Roesch, F.A., Jr.; Green, E.J.; Scott, C.T. 1989. New compatible estimators for survivor growth and ingrowth from remeasured horizontal point samples. Forest Science. 35(2): 281-293.

Scott, C.T. 1979. Northeastern forest survey board-foot volume equations. Res. Note NE-271. U.S. Department of Agriculture, Forest Service. 3 p.

Teck, R.M.; Hilt, D.E. 1991. Individual tree-diameter growth model for the Northeastern United States. Res. Pap. NE 649. U.S. Department of Agriculture, Forest Service. 10 p.

Van Deusen, P.C.; Dell, T.R.; Thomas, C.E. 1986. Volume growth estimation from permanent horizontal points. Forest Science. 32(2): 415-422.

NEURAL NETWORKS VS. MULTIPLE LINEAR REGRESSION FOR ESTIMATING PREVIOUS DIAMETER

Susan L. King[1]

Abstract—A neural network is a nonparametric statistical modeling procedure known for its capacity to process nonlinear relationships. For estimating the previous diameter of a tree, the exact functional relationship between the response variable and the independent variables is unknown. The relationship is most likely nonlinear. Multiple linear regression was used to develop a model for estimating the previous diameter of trees in West Virginia. The data were split into a model data set with 8,723 observations and a validation data set with 7,951 observations. The dependent variable was either basal-area increment or diameter increment. Two different sets of independent variables were evaluated. The data were divided into six species groups based on the rank of the average diameter growth of the species. Basal-area increment was a superior dependent variable for the multiple linear regression model. Basal-area increment is a nonlinear transformation of the diameter increment. It was thought that neural networks with its capacity to capture nonlinear relationships might provide an equivalent or superior solution with the diameter increment response variable as opposed to the basal-area increment response variable. All of the basal-area increment models had a higher R^2 than their diameter increment counterparts. Neither technique was superior in all cases. Other issues such as the stopping criteria, initial weight selection, the optimal number of hidden nodes, and the optimal number of hidden layers in a neural network are also discussed.

INTRODUCTION

The goal of a neural network is to mathematically model the brain and to capture its pattern recognition capabilities. Humans are more efficient at processing pattern information such as speech and visual images than any machine, whereas computers are extremely fast at processing information that can be formulated into a sequence of instructions.

A neural network may provide a superior solution over a traditional statistical approach for certain classes of problems (Burke, 1991). These classes include problems in which the distributions are unknown and possibly nonlinear, where outliers may exist, and where noise is present in the data. These are common conditions in forest inventory data. This paper investigates whether neural networks provide improved estimates over the traditional statistical modeling procedure of multiple linear regression for estimating diameter of a tree at an earlier time period.

In multiple linear regression, the relationship between independent and dependent variables is assumed to be linear and interactions among the independent variables must be specified in advance by the user. In neural networks, there is no assumed relationship between the independent and dependent variables. The relationship between independent and dependent variables and the interactions among the independent variables are learned through an iterative process. Neural networks require no assumptions about the distributions, mean, or correlation of the errors.

PROBLEM

The Northeastern Forest Inventory and Analysis (NEFIA) unit of the USDA Forest Service currently uses trees measured at current and previous inventories to develop multiple linear regression equations to estimate previous diameter for those trees not recorded at the previous inventory. The impetus for this study was to develop an improved procedure to estimate a previous diameter for ongrowth and nongrowth trees on plots sampled by variable radius plot sampling. However, the models built can be used whenever a previous diameter is required and the appropriate dependent variables are available. For West Virginia, King and Arner (1998) developed a new procedure to estimate a previous diameter. They used six groupings based on the rank of the average diameter growth for a species. Also in their study, many different independent variables and combinations of independent variables were evaluated. To investigate whether neural networks can provide a better estimate of previous diameter, the best models from the King and Arner study were selected for comparison.

DATA

The data for this project came from 1,965 remeasured forest inventory plots in West Virginia. The trees were initially measured in 1975 and then remeasured in 1988. Only trees larger than 5 inches in diameter at both time periods were included. The data were randomly split into a model building data set with 8,723 observations and a validation data set with 7,951 observations. Before splitting the data, they were grouped into six species groups. There were two sets of independent variables chosen for comparison. The first set of variables are those currently used by NEFIA to estimate the previous diameter. These variables are:

DBH2 = tree diameter at time period 2;

TRCLS2 = tree class at time period 2, a measure of tree quality;

[1] Operations Research Analyst, USDA Forest Service, 100 Matsonford Road, 5 Radnor Corp Ctr., Ste. 200, Radnor, PA 19087-4585.

Citation for proceedings: Stringer, Jeffrey W.; Loftis, David L., eds. 1999. Proceedings, 12th central hardwood forest conference; 1999 February 28-March 1-2; Lexington, KY. Gen. Tech. Rep. SRS-24. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 293 p. [Peer-reviewed paper].

CRNCLS2 = crown class at time 2, a measure of crown position in the canopy;

CRATIO2 = crown ratio at time 2, the proportion of a tree with a live crown;

CRCC2 = CRATIO2 / CRNCLS2;

DCR2 = DBH2 • CRATIO.

The analysis by King and Arner showed that the addition of the variable BAL2 improved the results. BAL2 is the sum of the basal areas of the trees on a plot larger than the subject tree. It is a measure of the competition for light. The addition of the variable BAL2 to the first set of independent variables formed the second set of independent variables. Two response variables were investigated: diameter increment (DI) and basal area increment (BAI).

$$DI = \frac{(DBH2 - DBH1)}{N} \quad (1)$$

or

$$BAI = \frac{K \cdot (DBH2^2 - DBH1^2)}{N} \quad (2)$$

where:

N = number of years between measurements on the plot;

DBH1 = tree diameter at time period 1;

DBH2 = tree diameter at time period 2;

K = 0.005454154, a conversion factor from diameter in inches to basal area in square feet.

Annual increment was used to account for variation in the measurement period among the plots. Basal area and dbh are related by a transformation. Logic would suggest that BAI would be a better response variable than DI. There is not a one-to-one mapping between DI and BAI. A poletimber and a sawtimber tree may have the same DI, but different diameters at both measurement periods. BAI captures the differences in the size of the trees. Thus, 24 models were compared for both multiple linear regression and neural networks.

NEURAL NETWORKS

The type of a neural network chosen for this study is a feedforward backward propagation network (Figure 1). The network consists of three layers: the input layer, hidden layer, and output layer. The layers consist of processing units called nodes. Arcs connect the layers. Each arc has a weight which represents the strength of the connection. The goal of a neural network is to find the best estimate of the weights.

In the input layer, the number of nodes corresponds to the number of independent variables. In the hidden layer, not only is the number of nodes variable, but also there may be more than one hidden layer. Only one hidden layer is shown in Figure 1. In general, only one hidden layer is required. The third layer is the output layer. The number of nodes in this layer corresponds to the number of dependent variables. A backpropagation network has no cycles. All of the arcs move from left to right.

Each observation in the data set forms an input pattern, p. An observation is called an exemplar in neural networks. A linear combination of the input patterns and the weights is formed at each hidden node. This defines a plane in N - 1 dimensional space, where N is the number of input nodes. The hyperplane passes through the origin unless a bias weight is added to the hidden node. Mathematically, this process is represented as:

$$net_j^p = w_j + \sum_{i=1}^{I} w_{ij} x_i^p \quad \text{for j=1,...,J} \quad (3)$$

where:

w_j = the bias term for hidden unit j;

w_{ij} = the weight from input node i to hidden node j;

x_i^p = i^{th} component of the p^{th} exemplar;

I = number of input nodes;

J = number of hidden nodes.

By creating a dummy node with a fixed input value of 1 or -1, the bias can be written as a weight, $w_{I+1,j}$. Thus, equation (3) becomes:

$$net_j^p = \sum_{i=1}^{I+1} w_{ij} x_i^p \quad \text{for j=1,...,J.} \quad (4)$$

A squashing or activation function is applied to net_j^p at each hidden node j. This function introduces nonlinearities into the network. Common squashing functions are the logistic and hyperbolic tangent function. Applying the squashing function to net_j^p yields:

$$Y_j^p = f(net_j^p) \quad \text{for j=1,...,J.} \quad (5)$$

The output from each of the j nodes at the hidden layer becomes the input to the k output nodes.

A linear combination of the output from the hidden nodes and the weights, v_{jk}, is formed. As before, a bias term is added. Mathematically, this may be expressed as:

$$net_k^p = \sum_{j=1}^{J+1} v_{jk} Y_j^p \quad \text{for k=1,...,K} \quad (6)$$

where:

v_{jk} = the weight from hidden node j to output node k.

A squashing function is applied to net_k^p to obtain the predicted output:

$$o_k^p = f(net_k^p) \quad \text{for k=1,...,K.} \quad (7)$$

Backpropagation is a supervised procedure. That is, it requires an observed dependent variable, t_k^p. The estimated value, o_k^p, is compared with t_k^p, to determine if they are close. One measure of closeness is the sum of the

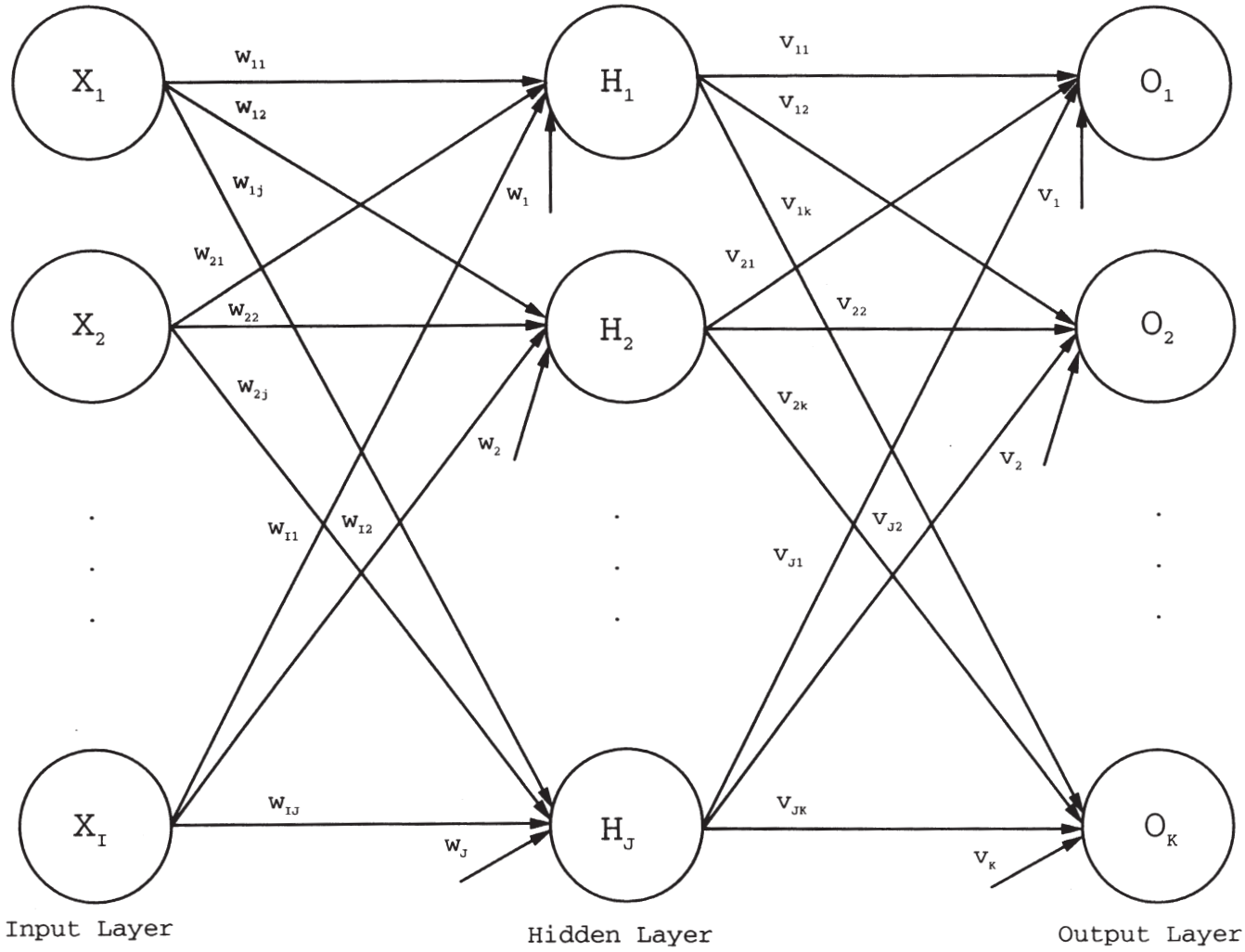

Figure 1—Backpropagation network. The network is fully connected. There is an arc from every node to a node in the next layer.

squared differences between t_k^p and o_k^p. It is used frequently because the derivatives are easy to compute. Thus, the objective function is:

$$\text{Min} \sum_p \sum_{k=1}^{K} (o_k^p - t_k^p)^2 . \quad (8)$$

The objective function for neural networks in equation (8) appears to be similar to that in multiple linear regression. Both techniques minimize the sum of the squared differences between the observed and the predicted values. The variables in both procedures are the weights. However, the two procedures are quite different. The predicted values are different functions of the weights. The number of weights in multiple linear regression depends on the number of input variables, whereas the number of weights in neural networks depends on the number of: input variables, hidden nodes, hidden layers, and output nodes. The objective function in both the neural network and multiple linear regression is an unconstrained minimization problem. The special structure in the multiple linear regression problem allows for the optimal set of weights to be found through solving a system of normal equations. Iterative techniques are used to find the optimal set of weights for a neural network. Each iteration is considered a training period. By updating the weights, the neural network is said to be learning.

Many techniques are available for solving unconstrained minimization problems. These techniques include gradient descent, the quasi-Newton techniques of conjugate gradients and Davidon-Fletcher-Powell, the modified Newton technique of Levenberg and Marquardt, stiff differential equations, genetic optimization, and simulated annealing. Historically, gradient descent has been applied by the neural network community to solve equation (8). In fact the name backpropagation refers to the process of applying the chain rule of calculus to compute the error gradient for each weight in the network. The error gradient is used in updating the weights in gradient descent. The error is said to be propagated backwards. Gradient descent may be advantageous if the problem is implemented on a parallel computer. However, most problems are implemented on a serial computer, and gradient descent on these machines has been abandoned by the optimization community in favor of more sophisticated techniques. The

difficulties of gradient descent are well documented. Sarle (1994), Masters (1995), Bishop (1995), and Bazarra, Sherali and Shetty (1993) all discuss the limitations of gradient descent and are excellent references on the conjugate gradient method, the Davidon-Fletcher-Powell, and the Levenberg-Marquardt algorithm. Kollias and Anastassiou (1988) discuss applying the Levenberg-Marquardt algorithm to neural networks. Owens and Filken (1989) saw the similarity between a system of stiff differential equations and the gradient descent approach. They claim that stiff differential equations provide a more rapid and accurate convergence than either gradient descent or conjugate gradient methods. Hassoun (1995) provides an introduction to simulated annealing and genetic optimization for neural networks. Masters (1995) is another good introductory reference for simulated annealing in neural networks. Another text by Masters (1993) provides introductory material on both topics.

In selecting the technique to solve equation (8), the number of weights must be taken into account. Sarle (1994) recommends using the Levenberg-Marquardt algorithm for networks with tens of weights, the Davidon-Fletcher-Powell algorithm for networks with hundreds of weights, and the conjugate gradient procedure for large problems with thousands of weights. Most of the 24 subproblems were solved using the Levenberg-Marquardt algorithm. These subproblems had only tens of weights. The other two techniques were tried on a few subproblems, but the value of their objective functions was larger. Gradient descent was also tried on several of the subproblems. It proved to be a superior technique for only the response variable DI in the first subgroup. The gradient descent algorithm was user written in SAS/IML(SAS Institute Inc., 1989). The other three algorithms are part of PROC NLP (SAS Institute Inc., 1997) in SAS/OR. SAS had a Beta release macro that was available in Release 6.10. This macro was modified and used in this study. Updated SAS neural network macros with a GUI interface are now part of the SAS Data Mining Solution.

Several other issues relating to the implementation of the neural network need to be discussed. First, the activation function must be selected. The purpose of an activation function is to induce nonlinearity into the network through a nonlinear transformation. With a linear function, the output is a weighted sum of the inputs. A squashing function is an activation function that maps any real input into a bounded range, usually between 0 and 1 or between -1 and 1. The two most common squashing functions are the logistic function and the hyperbolic tangent function. Other functions may be used so long as they are differentiable. Smooth activation functions decrease the training time. Kalman and Kwasny (1992) argue that the hyperbolic tangent function was the best of the sigmoidal functions. Bishop (1995) states that the hyperbolic tangent function often increases algorithm convergence over a logistic function. The choice of an activation function may be different at the output node. The range of the activation function at the output node should correspond to the range of the dependent variable. A categorical dependent variable would have a different activation function than a continuous output variable. The output data for the diameter increment problem were scaled. The hyperbolic tangent function was used at both the hidden and output layers. The logistic function was tried on some of the subproblems and it did not provide a superior solution.

Second, there are two different philosophies concerning the scaling of the independent and dependent variables. Some believe that it is not necessary to scale the independent or dependent variables. The size of the weights will make any necessary adjustments. In the diameter increment problem, the input and output variables were scaled to values that correspond to the range of the squashing function. Because the hyperbolic tangent function is used as the squashing function, the continuous variables are scaled between -0.9 and 0.9. The endpoints, -1 and 1, in the range are not used because they correspond to the inputs $-\infty$ and $+\infty$, respectively. The class variables are first broken into indicator variables and then scaled like the continuous variables. The lowest value of the indicator variable corresponds to -0.9 and the largest value corresponds to 0.9.

Third, the selection of the initial weights is a major issue. None of these techniques guarantee a global minimum. The choice of initial weights can influence the quality of a local solution. The initial weights in this project were selected by random numbers. There are many heuristic procedures for selecting the initial weights. One procedure by Piovoso and Owens (1991) was tried on several subproblems. In this procedure, the weights between the input layer and the hidden layer are found by principal component analysis. The weights between the hidden layer and the output layer are found by multiple linear regression. In this project, after trying several different seeds, a random number generator always provided a set of weights that found a lower value to the objective function than the Piovoso and Owens procedure.

Fourth, the number of layers must be selected. Hornik, Stinchcombe, and White (1989) showed that a neural network with one hidden layer and an arbitrary squashing function can approximate most functions. In practice, the need for a second hidden layer occurs when a piecewise-continuous function must be approximated. This condition is not present in the diameter or basal-area growth problem, so only one hidden layer was used.

Fifth, there are no equations or formulas for selecting the optimal number of hidden nodes. Each situation is different. Too many hidden units cause an inability to generalize and the data are overfitted. Similarly, too few hidden nodes will cause an underfitting of the data. An underfitted or an overfitted model does not generalize well, that is, predict accurate dependent or output variables from a new set of independent or input variables. One way to control generalization is through the selection of the number of hidden units and their connections. This is model selection. The simplest model selection technique is to determine the optimal number of hidden nodes through experimentation. Starting with one or two nodes, a solution is found. Another node is added and the problem is reoptimized. This

process is repeated until the objective function value begins to increase. The number of hidden nodes corresponding to the minimal objective function value is optimal. This procedure was used in the diameter increment model. In most of the 24 problems, two hidden nodes were optimal. An alternative is to start with a large number of hidden nodes and gradually remove complete hidden units or only remove selected connections. This is pruning. Reed (1993) describes several pruning procedures.

Sixth, in any iterative algorithm, a major issue is when to stop. The error in the model data set monatonically decreases as a function of the iteration number. The error in the validation data set decreases and then increases as the neural network starts to overfit. The algorithm is terminated when the validation data set reaches its minimum. This procedure is called stopped training. Regularization procedures such as stopped training improve generalization by controlling the size of the weights. Other regularization procedures such as weight decay, training with noise, and Bayesian estimation are described by Bishop (1995). Stopped training was used in this project.

COMPARISON STATISTICS

The statistics used to compare multiple linear regression with neural networks for both the model and the validation data set are R^2, the mean of the squared errors (MSE), the mean of the absolute errors (MAE), and the mean of the arithmetic errors (ME). The ME indicates bias, whereas the MSE and the MAE both indicate precision as well as bias. The MAE is more robust and less sensitive to outliers than the MSE.

RESULTS AND DISCUSSION

A comparison of the results between neural networks and multiple linear regression for the ranked mean species groups is presented in Tables 1 and 2. The overall results were obtained by combining the response values and the predicted response values for each of the six species groups and forming one group. The appropriate statistics are then calculated. Because the goal of a model is generalization, the results for the validation data set are more important than those for the model data set.

For the BAI models, all of the results are expressed as DI using the translation:

$$\hat{DI}(BA) = \frac{\sqrt{\frac{(BA1 + N \cdot \hat{BAI})}{K}} - DBH1}{N} \qquad (9)$$

where:

BA1 = tree basal area at time period 1, $K \bullet DBH1^2$;

$\hat{BAI}$ = predicted basal-area increment;

N = number of years between measurements of the tree;

All of the statistics use $\hat{DI}(BA)$ as the predicted diameter increment for the basal-area models.

From the overall results for the response variable DI, neural networks was superior. It had a slightly higher R^2, a slightly lower MAE, and a slightly lower MSE. For the individual species groups for the response variable DI, neural networks was superior for the NEFIA variables, but the results were mixed for addition of BAL2. Still, neural networks predominates. The ME was larger for the neural network model as expected. An assumption of multiple linear regression is that the expected value of the errors is zero. Neural networks, on the other hand, is a nonparametric procedure and the mean of the arithmetic error will not necessarily be zero. The addition of the variable BAL2 improved the results for both neural networks and multiple linear regression. For the response variable BAI, multiple linear regression more frequently provided a superior solution than neural networks as indicated by the R^2, MAE, and MSE.

The software used in this study was the first beta version of SAS's neural network macros. Later beta versions were significantly enhanced. A subset of the 24 subproblems was selected to access the impact of the new software on the diameter increment prediction problem. The subset had BAI as the response variable and NEFIA + BAL2 as the independent variables. Neural networks had the most difficulty with this subset. The results are in Table 3. With the new macros, neural networks is the winner for the model data set. The R^2 is higher for all species groups, and the MAE and MSE are lower for most of the species groups. However, the new macros did not improve the results for the validation data set. Only the three smallest species groups have a higher R^2 and a lower MAE in the validation data set as compared with the results in Table 2. Neural networks performed slightly worse for the third and sixth species groups. The new macros did not significantly alter the results; and so, it was decided not to pursue modeling the remaining subgroups. Also other independent variables from the King and Arner (1998) study were tried on this subset of the 24 problems. They did not improve the results.

Neural networks does not always outperform multiple linear regression. This conclusion was also reached by Desai and Bharati (1998). They found that for predicting excess returns on large stocks, neural networks outperformed multiple linear regression in periods of high volatility. Otherwise, multiple linear regression was superior. These results parallel a study by Markham and Rakes (1998). Using computer generated data, Markam and Rakes conclude that there is a significant interaction between sample size and variance. Multiple linear regression performs better for low-variance problems and neural networks performs better for high-variance problems. The results are mixed and dependent on sample size for medium-variance problems. Neural networks was superior for large sample sizes, and multiple linear regression was superior for small sample sizes. One explanation of why neural networks did not outperform multiple linear regression in this project is that there is not enough

Table 1—Comparison statistics for neural networks and mutiple linear regression for NEFIA variables

Sub-group	No. of trees	NN	REG	NN	REG	NN	REG	NN	REG
		----------R^2----------		----------ME----------		--------MAE--------		---------MSE---------	
				DI response variable and model data set					
1	257	0.0463	0.0449	-0.0051	0.0000	0.0334	0.0322	0.0017	0.0017
2	947	0.1099	0.0846	0.0004	0.0000	0.0409	0.0415	0.0027	0.0028
3	2489	0.1649	0.1568	0.0016	0.0000	0.0430	0.0434	0.0031	0.0032
4	2371	0.2359	0.2011	0.0019	0.0000	0.0541	0.0556	0.0049	0.0051
5	783	0.2449	0.2193	0.0004	0.0000	0.0557	0.0564	0.0050	0.0052
6	1876	0.2440	0.2116	0.0050	0.0000	0.0664	0.0689	0.0074	0.0077
All	8723	0.3731	0.3512	0.0020	0.0000	0.0517	0.0528	0.0046	0.0048
				DI response variable and validation data set					
1	97	0.0256	0.0011	-0.0020	0.0018	0.0335	0.0336	0.0017	0.0017
2	616	0.1106	0.0787	0.0011	0.0024	0.0419	0.0425	0.0029	0.0030
3	2417	0.1973	0.1832	0.0034	0.0001	0.0436	0.0436	0.0031	0.0031
4	2448	0.2262	0.1606	0.0013	0.0016	0.0544	0.0561	0.0049	0.0054
5	509	0.1184	0.1148	0.0105	0.0043	0.0620	0.0598	0.0062	0.0062
6	1864	0.2334	0.2015	0.0044	-0.0021	0.0686	0.0700	0.0077	0.0081
All	7951	0.3521	0.3216	0.0032	0.0005	0.0537	0.0545	0.0049	0.0051
				BAI response variable and model data set					
1	257	0.2481	0.2349	-0.0011	0.0011	0.0293	0.0288	0.0013	0.0013
2	947	0.2806	0.3155	0.0029	0.0004	0.0366	0.0359	0.0022	0.0021
3	2489	0.4005	0.3704	-0.0003	0.0012	0.0366	0.0374	0.0022	0.0024
4	2371	0.4679	0.4706	0.0007	0.0010	0.0456	0.0452	0.0034	0.0034
5	783	0.4371	0.4734	-0.0044	0.0021	0.0491	0.0464	0.0037	0.0035
6	1876	0.4353	0.5127	-0.0181	0.0022	0.0607	0.0540	0.0055	0.0048
All	8723	0.5418	0.5647	-0.0039	0.0013	0.0451	0.0435	0.0034	0.0032
				BAI response variable and validation data set					
1	97	0.1311	0.2036	0.0097	0.0022	0.0301	0.0300	0.0015	0.0014
2	616	0.3248	0.3153	-0.0043	0.0029	0.0374	0.0368	0.0022	0.0022
3	2417	0.3848	0.3921	0.0036	0.0168	0.0373	0.0373	0.0023	0.0023
4	2448	0.4539	0.4495	0.0018	0.0018	0.0455	0.0455	0.0035	0.0035
5	509	0.3751	0.4008	0.0051	0.0027	0.0508	0.0497	0.0044	0.0042
6	1864	0.4221	0.5143	0.0110	0.0014	0.0594	0.0551	0.0058	0.0049
All	7951	0.5222	0.5518	0.0044	0.0018	0.0458	0.0447	0.0036	0.0034

variability in the data. Another explanation is that the relationship between BAI and the independent variables is linear. In this situation, a neural network can not outperform a linear model.

CONCLUSIONS

Neither neural networks nor the multiple linear regression significantly outperformed the other technique. Neural networks followed the same trend as multiple linear regression. All of the statistics indicate significant improvement by using the response variable BAI instead of DI. The addition of the variable BAL2 also improved slightly both models as indicated by the statistics. King and Arner (1998) found that the addition or substitution of other variables had little impact on the model.

There is contradictory information in books and journals on neural networks. There is no consensus on the selection of the initial weights, model selection, regularization, and scaling of input and output variables. An effort has been made to try new architectures on a subset of the 24 models as they become available in SAS.

Table 2—Comparison statistics for neural networks and multiple linear regression for NEFIA variables + BAL2

Sub-group	No. of trees	NN	REG	NN	REG	NN	REG	NN	REG
		----------R^2----------		----------ME----------		--------MAE--------		---------MSE---------	
				DI response variable and model data set					
1	257	0.0702	0.0671	-0.0003	0.0000	0.0323	0.0324	0.0016	0.0016
2	947	0.1476	0.1200	0.0002	0.0000	0.0397	0.0405	0.0026	0.0026
3	2489	0.1933	0.1925	0.0044	0.0000	0.0419	0.0425	0.0030	0.0030
4	2371	0.2649	0.2311	0.0011	0.0000	0.0529	0.0548	0.0469	0.0049
5	783	0.2696	0.2628	-0.0026	0.0000	0.0553	0.0550	0.0048	0.0049
6	1876	0.3198	0.2882	0.0017	0.0000	0.0628	0.0652	0.0067	0.0070
All	8723	0.4097	0.3907	0.0017	0.0000	0.0501	0.0513	0.0043	0.0045
				DI response variable and validation data set					
1	97	0.0357	0.0515	0.0012	0.0014	0.0334	0.0330	0.0016	0.0016
2	616	0.1355	0.1231	0.0030	0.0010	0.0412	0.0413	0.0028	0.0028
3	2417	0.2076	0.2186	0.0092	-0.0001	0.0436	0.0426	0.0030	0.0030
4	2448	0.2241	0.1964	0.0015	0.0009	0.0547	0.0552	0.0049	0.0051
5	509	0.1578	0.1590	0.0064	0.0023	0.0594	0.0585	0.0059	0.0059
6	1864	0.2895	0.2672	0.0063	-0.0008	0.0664	0.0671	0.0072	0.0074
All	7951	0.3739	0.3612	0.0054	0.0003	0.0531	0.0530	0.0047	0.0048
				BAI response variable and model data set					
1	257	0.2426	0.2549	-0.0037	0.0012	0.0300	0.0287	0.0013	0.0013
2	947	0.3350	0.3391	0.0020	0.0005	0.0351	0.0352	0.0020	0.0020
3	2489	0.4005	0.3876	-0.0003	0.0013	0.0366	0.0369	0.0022	0.0023
4	2371	0.4811	0.4782	0.0006	0.0012	0.0451	0.0452	0.0033	0.0033
5	783	0.4895	0.4914	0.0005	0.0023	0.0457	0.0455	0.0034	0.0034
6	1876	0.5282	0.5355	-0.0120	0.0027	0.0548	0.0528	0.0046	0.0046
All	8723	0.5782	0.5782	-0.0023	0.0016	0.0433	0.0429	0.0031	0.0031
				BAI response variable and validation data set					
1	97	0.1623	0.2342	0.0093	0.0018	0.0296	0.0296	0.0014	0.0013
2	616	0.3599	0.3449	-0.0012	0.0016	0.0357	0.0359	0.0021	0.0021
3	2417	0.3848	0.4048	0.0036	0.0018	0.0373	0.0372	0.0023	0.0023
4	2448	0.4658	0.4619	0.0003	0.0013	0.0454	0.0454	0.0034	0.0034
5	509	0.4098	0.4161	0.0033	0.0019	0.0495	0.0494	0.0041	0.0041
6	1864	0.4655	0.5294	0.0171	0.0026	0.0566	0.0546	0.0054	0.0048
All	7951	0.5416	0.5634	0.0054	0.0019	0.0449	0.0444	0.0035	0.0033

This work represents another step in evaluating the power of neural networks. Finding a ‘good’ set of initial weights can be a time consuming process, and some thought must be given to the frequency of use and the required precision of the final model before abandoning traditional techniques.

Table 3—Comparison statistics for neural networks and multiple linear regression for DEFIA + BAL2 variables using an updated version of SAS macros

Sub-group	No. of trees	NN	REG	NN	REG	NN	REG	NN	REG
		----------R^2----------		----------ME----------		--------MAE--------		---------MSE---------	
				BAI response variable and response data set					
1	257	0.4828	0.2549	0.0010	0.0012	0.0229	0.0287	0.0009	0.0013
2	947	0.4104	0.3391	0.0022	0.0005	0.0330	0.0352	0.0018	0.0020
3	2489	0.4022	0.3876	0.0013	0.0013	0.0363	0.0369	0.0022	0.0023
4	2371	0.5030	0.4782	0.0018	0.0012	0.0436	0.0452	0.0032	0.0033
5	783	0.4932	0.4914	-0.0005	0.0023	0.0459	0.0455	0.0034	0.0034
6	1876	0.5634	0.5355	-0.0012	0.0027	0.0513	0.0528	0.0043	0.0046
All	8723	0.5991	0.5782	0.0008	0.0016	0.0416	0.0429	0.0029	0.0031
				BAI response variable and validation data set					
1	97	0.1964	0.2342	0.0094	0.0018	0.0278	0.0296	0.0014	0.0013
2	616	0.3665	0.3449	0.0030	0.0016	0.0349	0.0359	0.0021	0.0021
3	2417	0.3813	0.4048	0.0013	0.0018	0.0378	0.0372	0.0024	0.0023
4	2448	0.4659	0.4619	-0.0013	0.0013	0.0451	0.0454	0.0034	0.0034
5	509	0.4121	0.4161	0.0007	0.0019	0.0495	0.0494	0.0041	0.0041
6	1864	0.4619	0.5294	0.0247	0.0026	0.0555	0.0546	0.0054	0.0048
All	7951	0.5403	0.5634	0.0062	0.0019	0.0446	0.0444	0.0035	0.0033

REFERENCES

Bazaraa, Mokhtar S.; Sherali, Hanif D.; Shetty, C.M. 1993. Nonlinear programming. New York: John Wiley. 638 p.

Bishop, Christopher M. 1995. Neural networks for pattern recognition. New York: Oxford University Press. 482 p.

Burke, Laura Ignizio. 1991. Introduction to artificial neural systems for pattern recognition. Computers and Operations Research. 18(2): 211-220.

Desai, Vijay S.; Bharati, Rakesh. 1998. Annals of Operations Research. 78(1998): 127-163.

Hassoun, Mohamad H. 1995. Fundamentals of artificial neural networks. Camdridge, MA: The MIT Press. 511 p.

Hornik, Kurt; Stinchcombe, Maxwell; White, Halbert. 1989. Multilayer feedforward networks are universal approximators. Neural Networks. 2: 359-366.

Kalman, Barry L.; Kwasny, Stan C. 1992. Why tanh: choosing a sigmoidal function. Baltimore, MD: International Joint Conference on Neural Networks: 578-581.

King, Susan L.; Arner, Stanford L. [In press]. Estimating previous diameter for ingrowth trees on remeasured horizontal point samples. Proceedings 12th central hardwood forest conference; 1999 February 28 - March 3; Lexington, KY.

Kollias, Stefanos; Anastassiou, Dimitris. 1988. Adaptive training of multilayer neural networks using a least square estimation technique. In: Proceedings of IEEE international conference on neural networks I: 383-390.

Markham, Ina S.; Rakes, Terry R. 1998. The effect of sample size and variability of data on the comparative performance of artificial neural networks and regression. Computers and Operations Research. 25(4): 251-263.

Masters, Timothy. 1993. Pratical neural networks recipes in C++. San Diego: Academic Press. 493 p.

Masters, Timothy. 1995. Advanced algorithms for neural networks: a C++ sourcebook. New York: John Wiley. 431 p.

Owens, A.J.; Filkin, D.L. 1989. Efficient training of the back propagation network by solving a system of stiff ordinary differential equations. Washington, DC: International Joint Conference on Neural Networks II: 381-386.

Piovoso, Michael J.; Owens, Aaron J. 1991. Sensor data analysis using neural networks. San Padre Island, TX. Proceedings of the 4th international conference on chemical process control: 101-118.

Reed, Russell; Pruning Algorithms-A Survey. 1993. IEEE Transactions on Neural Networks. 4(5): 740-747.

Sarle, W.S. 1994. Neural network implementation in SAS software. Proceedings of the 19th annual SAS users group international conference. Cary, NC: SAS Institute Inc.: 1551-1573.

SAS Institute Inc. 1989. SAS/IML® Software: usage and reference. Version 6, 1st ed. Cary, NC: SAS Institute Inc. 501 p.

SAS Institute Inc. 1997. SAS/OR technical report: the NLP procedure. Cary NC: SAS Institute Inc. 146 p.