Using Informatics to Value Forest Stand Information
by
Matthew Pelkki and James Ringe

Abstract
Using an informatics approach to forest stand information, an innovative method of valuing inventory information about forest stands has been developed. The natural systems are modeled by individual-tree growth models within a discrete-state dynamic programming simulation system. Under varying state neighborhood sizes, the dynamic programming system returns different values for the financially optimal management regime. Using the state neighborhood interval size as a representation of the known precision around a particular stand variable, such as basal area, average diameter, or number of trees per acre, the marginal increase in the dynamic programming results can be used as a value for the increased precision in stand parameters. Since increased precision in mensuration data is often achieved at the expense of additional labor and resources, this methodology will assist with determining an economically efficient level of inventory expense. Results for shortleaf pine in the Central United States demonstrate the value of increased precision in stand information in terms of better management decisions and economic returns.

INTRODUCTION
The need to place a value on information collected about forest stands is well understood. Forest inventory typically samples a portion of a stand or forest management unit and provides information such as stand age, species composition, and stand density (basal area, volume, number of trees per unit area) that is used to guide management decisions. While the goal of increasing sampling intensity is to improve the accuracy of stand parameters, if the sampling technique is unbiased, it actually has the effect of increasing the precision of the parameter estimates. Of course, increasing sampling intensity results in a more expensive survey which can be measured through the cost of increasing the number of sample plots or using more precise measurement equipment (i.e., using a diameter tape rather than a Biltmore stick).

Typically, the value of more precise inventory information is related to the value of the current standing timber crop. It is important to recognize that greater precision in inventory data will lead to better management decisions over a long period of time. If stand objectives are measured in economic returns, then a more precise inventory will provide information to manage the stand to yield greater economic returns. The difficult question to answer is just how a greater level precision in stand information improves both current and future financial returns from a stand.

This study uses an informatics approach to exploring this problem that offers great promise in attaching a value to increasing levels of stand-level inventory information.

Informatics, as a scientific field of study developed during the expansion of information technology in the 1960's (Jones 1974). It has numerous applications in fields of data mining, artificial intelligence, knowledge encoding, multi-database management systems, healthcare and robotics (Hlavac et al 2001). Informatics can best be thought of as the study of the structure, behavior, and interactions within such a system.

The artificial system used to emulate a natural computational system is forward recursive dynamic programming (Dykstra 1984). The particular model used relies upon neighborhood storage technology for state variables (Brodie and Kao 1979, Arthaud and Klemperer 1985). Dynamic programming

2 Associate Professor and George H. Clippert Chair of Forest Resource Economics and Management, School of Forest Resources, Arkansas Forest Resources Center, University of Arkansas at Monticello, Monticello, AR 71656 and Professor, Department of Forestry, University of Kentucky, Lexington, KY 40546-0073.
requires discrete states to model a solution network. Since variables such as basal area or volume are typically continuous, state neighborhoods are created to act as “storage sites” in the dynamic programming network. For example, a particular dynamic programming formulation might use basal area (ft.$^2$/acre) and volume (ft.$^3$/acre) as state variables and define neighborhoods as "5 ft.$^2$" and "10 ft.$^3$". One particular state neighborhood in the network might be defined as ranging from 60 to 65 ft.$^2$ per acre in basal area and 1000 to 1010 ft.$^3$ per acre in volume. In essence, any actual stand state that fell within these parameters at a particular point in time would be considered to comparable in terms of future growth and value potential.

Previous studies (Valsta 1990, Pelkki 1997) have shown that as the state neighborhood interval is decreased, the objective function value the solution arrived at by dynamic programming improves until it converges on a global optimal value.

The informatics approach here is to assume that the size of the state neighborhood as a proxy for the precision of a particular stand parameter. Thus, the dynamic programming system simulates a natural computational system. As the quality of information available improves (as measured by greater precision or smaller state neighborhoods), the dynamic programming formulation is able to arrive at solutions (management prescriptions) that have higher economic returns.

METHODS

Multiple dynamic programming simulations were completed for shortleaf pine plantations in the central United States. The objective function was to maximize net present worth of an infinite series of rotations (soil expectation value). Six two-state combinations of four state variables (arithmetic mean diameter at breast height, basal area, number of trees per acre, and cubic foot volume per acre) were established, and for each combination, numerous dynamic programming simulations were made using different state neighborhood sizes in each run.

The resulting objective function values were compared to a global optimal value and the marginal gains from decreasing the size of state neighborhoods was calculated. These data were then fit using a linear model within spreadsheet software to create marginal cost curves for different combinations of state variables.

RESULTS

As state neighborhood sizes decreased in size, all dynamic programming simulations converged on a single, global optimal objective function value of $541.1/ha. In all cases, this objective function value was the result of a stand management schedule that was identical in terms of initial planting density, rotation length and thinning practices under all combinations of state variables using very small state neighborhoods.

Figures 1-6 show that, as state neighborhoods increase in size, the objective function value decreases by the amount shown.

Some general patterns appear in all the graphs. If the neighborhood size of one variable is small, then the marginal cost curve for the second variable is has little positive slope. This means that increasing the precision around a second state variable given that the first state variable is known precisely will add little to the objective function value. For example, in figure 1, if the state neighborhoods around number of trees per acre are small (4.9 or 12.4 trees per ha), then increasing the size of the basal area neighborhood has little impact on the objective function value.

Figure 1. Marginal cost of state neighborhood size
increases for BA-NT model.

Figure 2. Marginal cost of state neighborhood size increases for BA-DBH model.

Figure 3. Marginal cost of state neighborhood size increases for BA-VOL model.

Figure 4. Marginal cost of state neighborhood size increases for DBH-NT model.

Figure 5. Marginal cost of state neighborhood size increases for DBH-VOL model.
Figure 6. Marginal cost of state neighborhood size increases in NT-VOL model.

If the first state variable has a large neighborhood size, then increasing the size of the second state neighborhood results in a much larger reduction in the objective function value.

In some two-state network models, the sensitivity to one state variable is much stronger than another. This is especially evident in the BA-VOL (Figure 3) and NT-VOL (Figure 6) models. In both of these cases, when the volume neighborhood is smaller than 2.8 m$^3$/ha, the marginal cost curve is identical for all combinations of the second variable (either BA or NT).

Volume appears to be a good variable to use in combination with any of the other state variables. In all three of the models which include volume as a state variable, BA-VOL (Figure 3), DBH-VOL (Figure 5), and NT-VOL (Figure 6), the slope of the marginal cost curve is positive at the smallest tested neighborhood size for volume (0.35 m$^3$/ha). In the other models, the efficiency of the second variable seems to fall off much more quickly if volume is not included.

Also note that the linear fit causes some of the marginal cost curves to travel below the y-axis, or actually show an increase in the objective function value. Rather than using a linear fit, a simple linear regression, forcing the regression though the zero value on the y-axis would be better and will be incorporated in future models.

**DISCUSSION AND FUTURE WORK**

The marginal cost curves of course depend on economic parameters used in the dynamic programming simulation such as the cost of capital and stumpage prices. With changing economic parameters, it will be possible to develop very different marginal costs for decreasing precision of information about a stand. This is not necessarily a disadvantage, but rather, the informatics approach is flexible and can be adapted to incorporate sensitivity to different market conditions.

The informatics approach depends upon growth models to accurately reflect the growth response curve to a management regime. It also assumes, in this application, that the precision of data (as represented by the state neighborhood intervals) is held constant and known throughout the rotation. Altering the dynamic programming formulation, however, to permit variable sized state neighborhoods throughout the simulation is possible. The growth models employed in this simulation were individual-tree models (Miner et al 1988) and were used within their tested parameters. It is reasonable to assume that the results have a high degree of relative accuracy if not predictive accuracy.

A strong advantage to this approach is the ability to test the interaction of multiple variables to find the most efficient combination of stand variables upon which to base management decisions. In this case, it appears that number of trees and volume was the most efficient combination, followed closely by volume combined with either basal area or diameter.

This approach also values information based on the benefit it provides the decision-maker over an entire rotation, not simply the current stand value. This is of particular importance in younger stands, where management decisions are more typically made based on anticipated future returns rather than immediate revenues.

Future work in applying an informatics approach to valuing stand-level information will include the use of linear regression models rather than linear-fit models. This will eliminate the possibility of a negative intercept and provide additional statistical information on the quality of the marginal cost
curves. If necessary, nonlinear regression models may also be applied.

Finally, additional cover types, market conditions, and development of stage-dependent, variable-sized state neighborhoods will be employed to explore the usefulness of informatics to valuing stand information.

LITERATURE CITED


