This page is intentionally left blank.
ANALYSIS AND EVALUATION OF THE R-RATIO HEURISTIC FOR FOREST SCHEDULING

Silvana R. Nobre and Luiz C.E. Rodriguez

ABSTRACT

The R-ratio heuristic for forest scheduling (Rodriguez, 1994) solves large scale 0-1 integer programming versions of the Type I harvest scheduling problem (Johnson and Scheurman, 1977). This paper reviews its foundations and describes its evolution until now. Heuristic scheduling system patterns are used to represent the problem, to describe the evaluation function and to monitor the solution process. Its logic and elements are presented based on some artificial intelligence (AI) principles proposed to build heuristic processes. An AI approach based on intelligent agents provides the basis to analyze the R-ratio’s (i) escape strategy from local optima and (ii) its hybrid A*-greedy strategy to the solution search. AI concepts are also utilized to evaluate performance indicators of efficacy, measured by the proximity to the optimal solution of the non-integer linear programming relaxed version of the same problem, and efficiency, measured by penetrance and space complexity. For the test problems, the R-ratio strategy to escape from local optima proved efficacious given that several feasible solutions with objective function values below the range of 0.5% were obtained. And the R-ratio approach to find feasible solutions also proved efficient given its focus on a low cost strategy to select path searches.

KEYWORDS: Integer programming, harvest scheduling, forest management, heuristics, artificial intelligence.

INTRODUCTION

Johnson and Scheurman (1977) classified the various linear programming approaches to forest harvest scheduling into two basic types known as Model I and Model II. In Model I type problems, each decision variable defines one specific sequence of management interventions, while in Model II type problems each decision variable represents one single intervention.

Using Clutter and others (1983) representation of the basic Model I linear programming model, the following definitions are needed:

\[ N = \text{number of forest units} \]
\[ M = \text{number of management regimes} \]

\[ T = \text{number of management periods in the planning horizon} \]
\[ A_i = \text{area of forest unit } i \]
\[ X_{ik} = \text{area of forest unit } i \text{ assigned to management regime } k \]
\[ D_{ik} = \text{value (per unit of area) of management regime } k \text{ in forest unit } i \]
\[ V_{iktp} = \text{volume (per unit of area) of product } p \text{ harvested from forest unit } i \text{ in management period } t \text{ if management regime } k \text{ is used} \]
\[ V_{Min_p} \text{ and } V_{Max_p} = \text{minimum and maximum volumes of product } p \text{ in period } t \]

\[ i, k, p \text{ and } t = \text{identify forest units, management regimes, forest products and planning periods, respectively} \]
and the model becomes:

Maximize \[ Z = \sum_{i=1}^{N} \sum_{k=1}^{M} D_{ik} X_{ik} \] (1)

subject to:

\[ \sum_{k=1}^{M} X_{ik} \leq A_i \quad (i = 1,2,...,N) \] (2)

\[ \sum_{i=1}^{N} \sum_{k=1}^{M} v_{ikp} X_{ik} \geq V_{Minp} \quad (t = 1,2,...,T) \quad (p = 1,2,...,P) \] (3)

\[ \sum_{i=1}^{N} \sum_{k=1}^{M} v_{ikp} X_{ik} \leq V_{Maxp} \quad (t = 1,2,...,T) \quad (p = 1,2,...,P) \] (4)

The R-ratio heuristic for forest scheduling (Rodriguez, 1994) solves large scale 0-1 integer programming versions of the problem stated in sentences (1) to (4). The solution domain becomes \( X_{ik} \in \{0,1\} \), \( D_{ik} \) and \( v_{ikp} \) are restated as follows:

\( D_{ik} \) = total value of management regime \( k \) used in forest unit \( i \)

\( v_{ikp} \) = total volume of product \( p \) harvested from forest unit \( i \) in management period \( t \) if management regime \( k \) is used

and sentence (2) becomes:

\[ \sum_{k=1}^{M} X_{ik} = 1 \quad (i = 1,2,...,N) \] (5)

In real applications, such problems present tens or even hundreds of thousands of decision variables, and optimal solutions are hard, or even impossible, to obtain when the branch-and-bound algorithm and available computing resources are used.

This paper presents some attributes of the R-ratio heuristic for forest scheduling. Patterns are used to represent the problem, to describe the evaluation function and to monitor the solution process. In the following sections, the logic of the R-ratio heuristic and some of its elements are presented based on artificial intelligence principles proposed to build heuristic processes.

Patterns in heuristic scheduling systems

Scheduling involves assigning courses of action that consume various resources in limited supply for certain periods of time. The courses of action in forest harvest scheduling problems may be called management regimes that in turn are composed of activities or operations. Each activity requires certain amounts of specified resources and results in certain amounts of outputs.

Forest scheduling problems are often complicated by a large numbers of constraints relating activities to each other, resources to activities, and either resources or activities to desirable or undesirable consequences. Since these complicated relationships can make exact solutions for large scheduling problems very difficult, it might be helpful to simplify the problem.

Problem representation

First it is assumed that all desirable forest management regimes, defined along the planning horizon as combinations of different sequences of important forest activities, are known for each forest unit. The consumption of limited resources, and consequential outputs, are also well known for each activity in the management regime. Then, the problem reduces to the selection of one single management regime for each forest unit that results in the best scheduling value possible, the maintenance of desirable outcomes above minimum levels, and the consumption of limited resources bellow maximum levels. The idea behind the R-ratio heuristic is to provide an efficient and a successful search strategy. During the development of the R-ratio, one of the key issues was the definition of a useful evaluation function for the search.

Evaluation function

An evaluation function measures the goodness of changing the present solution. In other words, it returns a number describing the desirability (or lack thereof) of changing the present solution arrangement. The R-ratio heuristic starts with the optimal solution for the unconstrained version of the forest scheduling problem (a relaxed solution to the problem). In other words, the solution for a problem formed only by sentences (1) and (2) forms the basis \{ \( X_{ik} = 1 \mid D_{ik} \) is max \}.

Constraints are then reinserted, and desirability translates into the level of how well new solutions reduce minimally the objective function value with maximum reduction in the level of infeasibilities. The R-ratio heuristic uses an evaluation function that is a function of these two reductions. Specifically, the R-ratio heuristics evaluates the desirability of new solutions based in the following ratio:

\[ R = \frac{\Delta_{inf}^{of}_{i,k}}{\Delta_{obj}^{of}_{i,k}} \]

where:

\( \Delta_{inf}^{of}_{i,k} \): reduction in the level of infeasibilities if management regime \( k \) in management unit \( i \) becomes part of the solution; and
\( \Delta^{obj}_{i,k} \): reduction in the objective function value if management regime \( k \) in management unit \( i \) becomes part of the solution.

The level of infeasibility, used to obtain \( \Delta^{inf}_{i,k} \), is calculated as the sum of absolute deviations (SumAD) from the target values (below VMintp and above VManptp) defined in constraints (3) and (4). The value of \( \Delta^{obj}_{i,k} \) represents the variation in SumAD when a management regime is replaced in one forest unit. The larger the variation, higher is the chance of the management regime responsible for that variation to become part of the solution.

As a side effect, the replacement of regimes also induces a variation in the objective function value. That variation is represented by \( \Delta^{obj}_{i,k} \). It was defined as a reduction because it is more likely for the objective function value to decrease, after a change in the basis, than to increase, given that the objective function value initially reflects the optimum solution for the unconstrained formulation of the problem.

At each iteration, a pair of values \( \Delta^{inf}_{i,k} \) and \( \Delta^{obj}_{i,k} \) is calculated for each variable not in the basis, producing as many R-ratios as the number of non basic variables. Very efficient routines can be programmed to produce the R-ratios, with consecutive iterations showing declining processing times.

**Solution process**

As mentioned before, the R-ratio heuristic starts with an optimal solution for the forest scheduling problem formed by sentences (1) and (2). Once constraints (3) and (4) are reintroduced, and given the integer nature of the problem – with the necessary assignment of one single regime for each management unit – new solutions necessarily result from replacements in the set of previously chosen management regimes. Due to the replacement process established by the heuristic, three very distinctive sets of management regimes arise: (i) the set of chosen regimes; (ii) the set of non-selected regimes; and (iii) the set of discarded regimes.

In 1994, Rodriguez proposed the R-ratio heuristic with replacements drawn from the set of non-selected regimes only. Rodriguez and McTague suggested in 1997 a new replacement strategy, with periodic visits, during the heuristic process, to the set of discarded regimes. During these visits, the regime in the set of discarded regimes, contributing the most to reduce infeasibilities, replaces the previously selected regime to the respective management unit.

Two parameters were then proposed: the iteration number \( \alpha \) at which the heuristic visits the set of discarded regimes for the first time, and the \( \beta \) number of iterations between visits to the set of discarded regimes. After a certain number of iterations, and given a feasible solution is found, all regimes in the non-selected and discarded sets are finally evaluated in terms of their potentiality to maintain the solution feasible and to improve the objective function value. Once none is found, the process is terminated. And the set of chosen regimes is presented as the best solution.

Iterative techniques, like the one above described for the R-ratio technique, play an important role among all optimization procedures. The general step of an iterative procedure goes from a current solution \( i \) to next solution \( j \) and check whether one should stop there or perform another step. Usually, one approach builds an iterative procedure starting with a neighborhood \( N(i) \) defined for each feasible solution \( i \), and the next solution \( j \) is searched among the solutions in \( N(i) \).

The R-ratio technique, contrary to most iterative neighborhood search methods, does not disregard unfeasible solutions. In fact, it starts with the most probably unfeasible solution, the unconstrained optimum. During the search process, the R-ratio technique indirectly keeps track of all previously visited solutions by storing the replaced variables in a set called discarded set. Variables in this set may become part again of a future solution - feasible or not - in the next iteration.

In a certain way, the R-ratio method can be compared to Tabu Search (TS). TS is considered a neighborhood search method (Hertz, Taillard and de Werra 1995; Glover and Laguna, 1998). Similarly to a general TS procedure, the R-ratio procedure temporarily excludes a given group of prospective alternatives from the set of searchable strategies, by forbidding variables to become members of a new solution. On the other side, though, the R-ratio does not maintain fixed the size or structure of its tabu list, represented by the discarded set. And, in fact, the number of variables in the R-ratio discarded set augments as the number of iterations grows.

The tabu list and the discarded set may look similar in terms of constraining the search process, but they are built based on very different principles. Basically, the tabu list in a TS procedure is formed by feasible solutions in a subset of \( N(i) \) forbidden to be visited by rules expressed in a systematically updated guiding memory. In the R-ratio procedure, the discarded set contains variables, highly valuable in terms of the objective function value, but undesirable given the unfeasibility they generate if introduced in the selected set.
The R-ratio method seems capable to compete with other heuristic techniques, including TS procedures. This paper, based on some artificial intelligence principles, presents an evaluation of the heuristic solution process provided by the R-ratio method.

Artificial intelligence principles

Nilsson (1982) refers to artificial intelligence systems as systems where a database is affected by well defined operations or production rules subject to control strategies. The R-ratio (i) manages a database with three data sets (chosen, selected and discarded regimes), (ii) establishes a production rule where the regime with the largest R-ratio enters the chosen set replacing the previously selected regime for the same management unit, and (iii) the control strategy is the routine that calculates the R-ratio for all regimes, chooses the regime with the largest R-ratio value, operates the replacement and tests the termination condition (no infeasibilities and no way to improve the objective function value).

Rich and Night (1993) and Nilsson (1982) refer to the following concepts: state, movement and goal. State is the database configuration at a certain moment. The database state is altered when a rule is applied, generating a movement. And the goal is the most wanted state. The set of all possible states is the problem space. In a framework of states and movements, the problem solution is the sequence of movements that takes the control system from an initial state to the goal state.

In the current implementation of the R-ratio heuristic, a state is defined by how the main structure assigns regimes to the three sets of selected, non-selected and discarded regimes. A movement is made every time one of these assignments is modified by the control system. Each transfer of regimes between sets defines a new state. The goal is the state where infeasibilities have all been eliminated.

Nilsson (1982) distinguishes two major kinds of control strategies: irrevocable and tentative. In an irrevocable control regime, an applicable rule is selected and applied irrevocably without provision for reconsideration later. In a tentative control approach, an applicable rule is selected and applied, but provision is made to return later to this point in the computation to apply some other rule. The R-ratio heuristic strategy does not work either irrevocably or tentatively. Regimes replaced by better regimes are placed in the set of discarded regimes. This set is revisited periodically. When a previously discarded regime is selected again, the resulting state does not reflect exactly the conditions prevailing when that regime was considered before.

The same author also refers to commutative and decomposable systems, and how certain conditions must exist to provide a certain freedom in the order in which rules are applied to such systems. In a commutative system the rules can be organized into an arbitrary sequence and applied to the database to produce a result independent of order. Decomposable systems generate a database that can be decomposed or split into separate components, independently treatable generating similar results to those obtained when rules are applied to the whole database. The R-ratio heuristic is commutative because changes in the replacement order, of a given set or replacements, result in the same goal state. Decomposability is not allowed by the R-ratio heuristic approach, given that the evaluation function has to be applied to the complete database in order to find the best replacement.

Rich and Knight (1991) refer to the concept of a heuristic function to guide the search process, suggesting the path to be followed in a branching process. A good heuristic function offers low costs to the heuristic solution process. But there is a trade off in terms of the quality of the goal state, and the ideal cost will have to balance longer solution paths with direct solutions. The R-ratio is a very simple ratio of two easily calculated numbers used as a heuristic function to calculate the distance from the present state to the goal state. Although simple, the concept represents the original problem very properly.

Search strategies are usually classified as depth-first or breadth-first. Supposing an initially uninformed scheme to order the nodes in a search tree, the depth-first strategy will always select the deepest node to expand; meanwhile the breadth-first strategy expands nodes along the same level of depth in the search tree. The R-ratio starts with an informed scheme to select the regimes in the initial state. These are the ones with the highest objective function value. The initial state changes as regimes are replaced, one at a time. The search occurs in the non-selected set of regimes, which by turn decreases in size as the process continues. Therefore, it can be stated that the R-ratio uses a depth-first approach each time a replacement has to be done. Provided the set of non-selected regimes is always finite and that a path will always exist, it is guaranteed the R-ratio finds a path to a final state.

The intelligent agent concept (Russel and Norvig, 1995) can also be utilized to evaluate the goodness of the results provided by the R-ratio heuristic. An agent perceives its environment through sensors and acts upon the environment through effectors. Intelligent agents act rationally and try to maximize a certain performance measure. Implemented as
replacing regimes in the selected set. The R-ratio softbot sensor perceives the environment calculating and evaluating the possibilities around. The agent decides rationally, maximizes its performance choosing the regime in terms of consuming computational resources, given that there is no need to save information from previous states; and (iv) there is no guarantee that the agent will find the optimal solution, but given a final feasible solution is obtained, the search strategy will maintain the agent very close to it.

To overcome the limitations brought by the incomplete search strategy taken by the R-ratio agent, Rodriguez and McTague (1997) suggested multiple runs with different values for the $\alpha$ and $\beta$ parameters used in each run. Given the very low number of iterations taken by the R-ratio search strategy to find a final state on each run, the practical results obtained by their suggestion revealed very promising. After several runs, the best solution will be the one with the highest objective function value among the feasible final states.

Russel and Norvig (1995) define the state space of the problem as the set of all states reachable from the initial state by any sequence of actions. A successor function defines an action in terms of which state will be reached by carrying out the action in a particular state. A path in the state space is simply any sequence of actions leading from one state to another. A goal test or the termination test is needed by the softbot to determine if a certain state is the goal state. Finally, it may be the case that one solution is preferable to another, even though they both reach the goal. Therefore, a path cost function is a function that assigns a cost to a path.

In the R-ratio heuristic, the successor function rules the decision to visit the set of discarded regimes. It governs upon what set of regimes the replacement selection process will occur, and behaves depending on the values of $\alpha$ and $\beta$ which are defined before the softbot starts its search. The termination test is simply based in the amount of infeasibilities, which must be ideally zero. The path cost function is the R-ratio itself calculated for all non selected regimes.

For the R-ratio heuristic agent, it can be said: (i) The agent does not make a complete search, once heading in one direction it will not return to test other paths (a depth-first behavior); (ii) time spent by the agent is small and decreases as the search gets to an end; (iii) it is not demanding in terms of consuming computational resources, given that there is no need to save information from previous states; and (iv) there is no guarantee that the agent will find the optimal solution, but given a final feasible solution is obtained, the search strategy will maintain the agent very close to it.

For the R-ratio heuristic agent, it can be said: (i) The agent does not make a complete search, once heading in one direction it will not return to test other paths (a depth-first behavior); (ii) time spent by the agent is small and decreases as the search gets to an end; (iii) it is not demanding in terms of consuming computational resources, given that there is no need to save information from previous states; and (iv) there is no guarantee that the agent will find the optimal solution, but given a final feasible solution is obtained, the search strategy will maintain the agent very close to it.

For the R-ratio heuristic agent, it can be said: (i) The agent does not make a complete search, once heading in one direction it will not return to test other paths (a depth-first behavior); (ii) time spent by the agent is small and decreases as the search gets to an end; (iii) it is not demanding in terms of consuming computational resources, given that there is no need to save information from previous states; and (iv) there is no guarantee that the agent will find the optimal solution, but given a final feasible solution is obtained, the search strategy will maintain the agent very close to it.
“cost of reducing the objective function value of the present state”. Both approaches base their movements on proxies of total costs and benefits.

Therefore, it can be said that the R-ratio has some of the characteristics of a greedy search, when it cuts the search cost considerably, and also some of the characteristics of the A* search. It can also be said that the R-ratio heuristic, although developed intuitively with no support from the Artificial Intelligence Theory, has included several of the concepts introduced by two relevant AI references (mainly Nilsson, 1982; and Russel and Norvig, 1995).

**Performance evaluation**

Evaluations of the R-ratio heuristic presented in this paper are based on results obtained from tests with 9 different problems. The evaluation tried to offer answers to the four main questions proposed by Russel and Norvig (1995):

i. does the strategy guarantee a solution when there is one? (completeness);
ii. how long does it take to find a solution? (time complexity);
iii. how much memory does it need to perform the search? (space complexity); and
iv. does the strategy find the highest-quality solution when there are different solutions? (optimality).

**Efficacy** is measured by the proximity to the optimal solution of the non-integer linear programming. **Efficiency** is measured by means of the concepts penetrance and space complexity. The penetrance, $P$, according to Nilsson (1982) of a search is “the extent to which the search has focused toward a goal, rather than wandered off in irrelevant directions”. It is defined as

$$P = \frac{L}{T}$$

where

$L$ is the length of the path found to the goal (total number of iterations) and

$T$ is the total number of nodes generated during the search.

Adapted to the R-ratio search, $T$ is the total amount of regimes evaluated in the set of non-selected regimes. Considering that each visit to the set of non-selected regimes has to calculate the R-ratio for every regime in the set, the number of evaluations in each iteration becomes:

1$^{\text{st}}$ iteration $(N - U) - 1$
2$^{\text{nd}}$ iteration $(N - U) - 2$
3$^{\text{rd}}$ iteration $(N - U) - 3$

\[ \text{n}^{\text{th}} \text{ iteration } (N - U) - n \]

where

$N$ is the total number of possible management regimes in the problem (integer variables), and

$U$ is the total number of management units.

Also, let $L'$ be the number of iterations visiting the set of non-selected regimes:

$$L' = L - 1 - \frac{(L - \alpha)}{\beta}$$

And, then, the total amount of regimes evaluated in the set of non-selected regimes along the search becomes:

$$T = \sum_{n=1}^{L'}((N - U) - n) = (N - U)l' - \frac{L'(L' + 1)}{2}$$

Space complexity, $C$, evaluates the maximum number of states to be investigated by the agent. It is calculated multiplying the maximum number of iterations by the maximum quantity of possible states in each iteration. For the R-ratio search, the maximum number of iterations is the result of adding the total number of possible visits to the set of non-selected regimes plus the number of visits to the set of discarded regimes. In the set of non-selected regimes, the maximum number of visits the agent can make is equal the total number of regimes in the problem. The amount of visits to the set of discarded regimes depends on the values of $\alpha$ and $\beta$. Therefore, space complexity becomes:

$$C = (N + \frac{N - \alpha}{\beta} + 1)(N - U)$$

**RESULTS**

Table 1 presents some of the characteristics of the 9 problems used to evaluate the R-ratio forest scheduling heuristic search. Problems vary in terms of number of binary 0-1 decision variables, number of forest management units, number of periods in the planning horizon and periodical minimum volumes. Table 1 also presents three optimal objective function values, one for the continuous unconstrained, one for the continuous constrained, and one for the integer constrained versions of each problem. Reasonable amounts of time and computing resources were made available for the branch-and-bound algorithm to produce optimal solutions. For problems 5 to 9, though, the algorithm reached its limitations before finding an acceptable solution.

Several different strategies, varying the R-ratio $\alpha$ and $\beta$ parameters generated a reasonable large set of good solutions. Table 3 reveals that for the two problems with tighter
CONCLUSION

For the tested problems, the R-ratio strategy to escape from local optima produced feasible solutions in the high percentage of 95.9% to 76.2%, which is an interesting result given the complexity of the problems.

Table 1—Dimensions and characteristics of the test problems

<table>
<thead>
<tr>
<th>Problem Number</th>
<th>Number of Integer Decision Variables</th>
<th>Number of Forest Management Units</th>
<th>Planning Horizon (periods)</th>
<th>Minimum Volume Constraint (m³)</th>
<th>Maximum Objective Function Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unconstrained formulation</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>20</td>
<td>21</td>
<td>5,000</td>
<td>217,372.00</td>
</tr>
<tr>
<td>2</td>
<td>155</td>
<td>30</td>
<td>21</td>
<td>8,000</td>
<td>297,390.00</td>
</tr>
<tr>
<td>3</td>
<td>288</td>
<td>32</td>
<td>21</td>
<td>18,000</td>
<td>3,588,282.89</td>
</tr>
<tr>
<td>4</td>
<td>288</td>
<td>32</td>
<td>21</td>
<td>19,000</td>
<td>3,588,282.89</td>
</tr>
<tr>
<td>5</td>
<td>2,584</td>
<td>152</td>
<td>15</td>
<td>10,000,000</td>
<td>275,057,822.00</td>
</tr>
<tr>
<td>6</td>
<td>2,584</td>
<td>152</td>
<td>15</td>
<td>12,000,000</td>
<td>275,057,822.00</td>
</tr>
<tr>
<td>7</td>
<td>605</td>
<td>70</td>
<td>15</td>
<td>4,000,000</td>
<td>95,098,613.00</td>
</tr>
<tr>
<td>8</td>
<td>17,620</td>
<td>2,102</td>
<td>15</td>
<td>115,000,000</td>
<td>3,569,538,192.00</td>
</tr>
<tr>
<td>9</td>
<td>17,620</td>
<td>2,102</td>
<td>15</td>
<td>121,000,000</td>
<td>3,569,538,192.00</td>
</tr>
</tbody>
</table>
Table 2—Performance of the R-ratio search strategy - efficiency and optimality

<table>
<thead>
<tr>
<th>Problem number</th>
<th>Solutions generated</th>
<th>Iterations</th>
<th>% of the Best</th>
<th>Iterations</th>
<th>% of the Best</th>
<th>Iterations</th>
<th>% of the Best</th>
<th>% of the unconstrained</th>
<th>% of the (continuous)</th>
<th>% of the (integer)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>744</td>
<td>56</td>
<td>95.29</td>
<td>101</td>
<td>97.37</td>
<td>41</td>
<td>98.76</td>
<td>93.91</td>
<td>98.40</td>
<td>99.90</td>
</tr>
<tr>
<td>2</td>
<td>741</td>
<td>151</td>
<td>95.24</td>
<td>183</td>
<td>99.86</td>
<td>59</td>
<td>98.82</td>
<td>93.31</td>
<td>99.10</td>
<td>99.84</td>
</tr>
<tr>
<td>4</td>
<td>632</td>
<td>234</td>
<td>98.91</td>
<td>245</td>
<td>99.15</td>
<td>83</td>
<td>99.93</td>
<td>99.63</td>
<td>99.90</td>
<td>99.96</td>
</tr>
<tr>
<td>5</td>
<td>1607</td>
<td>2001</td>
<td>90.48</td>
<td>2456</td>
<td>95.73</td>
<td>1008</td>
<td>99.64</td>
<td>89.97</td>
<td>99.39</td>
<td>—</td>
</tr>
<tr>
<td>6</td>
<td>1001</td>
<td>1642</td>
<td>95.43</td>
<td>2564</td>
<td>96.06</td>
<td>887</td>
<td>99.93</td>
<td>91.77</td>
<td>99.71</td>
<td>—</td>
</tr>
<tr>
<td>7</td>
<td>6878</td>
<td>301</td>
<td>90.13</td>
<td>335</td>
<td>95.92</td>
<td>143</td>
<td>99.94</td>
<td>94.53</td>
<td>99.61</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>304</td>
<td>4201</td>
<td>99.28</td>
<td>5001</td>
<td>99.56</td>
<td>3599</td>
<td>99.95</td>
<td>97.81</td>
<td>99.84</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>252</td>
<td>5001</td>
<td>98.63</td>
<td>5001</td>
<td>98.63</td>
<td>3853</td>
<td>99.88</td>
<td>97.56</td>
<td>99.73</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 3—Performance of the R-ratio search strategy - penetrance

<table>
<thead>
<tr>
<th>Problem number</th>
<th>Solutions generated</th>
<th>#</th>
<th>%</th>
<th>Best solution (iterations)</th>
<th>#</th>
<th>%</th>
<th>Regime evaluations (expanded nodes)</th>
<th>Penetrance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>744</td>
<td>177</td>
<td>23.8</td>
<td>50</td>
<td>35</td>
<td>70.0</td>
<td>2,188</td>
<td>0.022857</td>
</tr>
<tr>
<td>2</td>
<td>741</td>
<td>321</td>
<td>43.3</td>
<td>99</td>
<td>64</td>
<td>64.6</td>
<td>5,921</td>
<td>0.016719</td>
</tr>
<tr>
<td>3</td>
<td>631</td>
<td>110</td>
<td>17.4</td>
<td>85</td>
<td>74</td>
<td>87.0</td>
<td>16,206</td>
<td>0.005245</td>
</tr>
<tr>
<td>4</td>
<td>632</td>
<td>26</td>
<td>4.1</td>
<td>102</td>
<td>84</td>
<td>82.4</td>
<td>17,942</td>
<td>0.005685</td>
</tr>
<tr>
<td>5</td>
<td>1,607</td>
<td>260</td>
<td>16.2</td>
<td>1,137</td>
<td>760</td>
<td>66.8</td>
<td>1,559,520</td>
<td>0.000729</td>
</tr>
<tr>
<td>6</td>
<td>1,001</td>
<td>74</td>
<td>7.4</td>
<td>918</td>
<td>614</td>
<td>66.9</td>
<td>1,304,750</td>
<td>0.000704</td>
</tr>
<tr>
<td>7</td>
<td>6,878</td>
<td>576</td>
<td>8.4</td>
<td>169</td>
<td>115</td>
<td>68.0</td>
<td>54,772</td>
<td>0.003085</td>
</tr>
<tr>
<td>8</td>
<td>304</td>
<td>35</td>
<td>11.5</td>
<td>3,680</td>
<td>3,097</td>
<td>84.2</td>
<td>43,267,682</td>
<td>0.000085</td>
</tr>
<tr>
<td>9</td>
<td>252</td>
<td>88</td>
<td>34.9</td>
<td>4,204</td>
<td>3,376</td>
<td>80.3</td>
<td>46,692,508</td>
<td>0.000090</td>
</tr>
</tbody>
</table>
range of 0.5% from the optimal non-integer solution. Considering the two artificial intelligence performance indicators used to analyze the R-ratio heuristic – penetrance and complexity – the R-ratio heuristic proved to be very efficient, given its focus on low cost strategies to select path searches. Also, given the obtained results, the R-ratio can be recommended as an efficient approach to find very good initial solutions for the integer version of the forest scheduling model I type problem, especially for the large scale problems usually found in real world applications.

**LITERATURE CITED**


This page is intentionally left blank.
IMPROVED SOLUTION TECHNIQUES FOR MULTIPERIOD AREA-BASED HARVEST SCHEDULING PROBLEMS

Juan Pablo Vielma\textsuperscript{1}, Alan T. Murray\textsuperscript{2}, David Ryan\textsuperscript{3}, Andres Weintraub\textsuperscript{4}

\textbf{ABSTRACT}

Area-based harvest scheduling models, where management decisions are made for relatively small units subject to a maximum harvest area restriction, are known to be very difficult to solve by exact techniques. Previous research has developed good approaches for solving small and medium sized forestry applications based on projecting the problem onto a cluster graph for which cliques can be applied. However, as multiple time periods become of interest, current approaches encounter difficulties preventing successful identification of optimal solutions. In this paper we present an approach for elasticizing timber demand constraints, which lends itself to an efficient solution technique. It is also possible using this approach to examine trade-offs between objective value performance and maintaining demand constraints.

\textbf{INTRODUCTION}

Mathematical modeling has been frequently used for harvest schedule planning. This has allowed several regulations and requirements to be incorporated in the planning process. These regulations are generally incorporated as restrictions to a Linear Integer Programming model and often make the problem more difficult to solve.

Regulations limiting spatial disturbances have led to constraints, typically known as maximum area restrictions, limiting the size of clear cut areas (Thompson et al. 1973, Murray 1999). Several models using these constraints have been proposed over the years, but the model known as the Area Restriction Model (ARM) has been shown to deliver the most profitable harvest schedules (Murray and Weintraub 2002). Unfortunately the ARM has proven to be very difficult to solve computationally. Although several heuristics to solve this model have been proposed (Hokans 1983, Lockwood and Moore 1993, Barrett et al. 1998, Clark et al. 1999, Richards and Gunn 2000, Boston and Bettinger 2001), exact methods have only recently been able to solve small and medium problem instances. One such method is that developed in Goycoolea et al. (2003), focusing on a strengthened formulation known as the Cluster Packing Problem. They were able to solve modest sized problems using a commercial integer programming solver for single period application instances. While solvability for multiple planning periods is possible, adding volume production restrictions creates significant complications for problem solution.

In this work we present an alternative way of structuring volume restrictions in order to restore most of the favorable properties of the single period Cluster Packing Problem. Application results are presented which demonstrate that near optimal solutions can be obtained quickly using the developed modeling approach.
HARVEST SCHEDULING WITH SPATIAL CONSTRAINTS

The harvest scheduling problem consists of selecting which areas of a forest will be harvested in different periods. Different types of requirements can be added to the generated harvested schedules. One environmental constraint that is generally enforced limits the contiguous area that can be harvested in any period. These constraints are generally known as maximum area restrictions (Thompson et al. 1973, Murray 1999).

We will assume that the forest is divided into sectors whose area is smaller than the maximum area that can be harvested contiguously and we will solve the harvest scheduling model known as Area Restriction Model (ARM). We will also assume a green-up time of one period. Finally we assume that each sector of the forest can only be harvested once during the planning horizon and that some kind of smoothing constraints over the volume of timber produced are desirable. Our base ARM formulation will be the Cluster Packing Problem developed in Goycoolea et al. (2003).

CLUSTER PACKING HARVEST SCHEDULING MODEL

The Cluster Packing Problem uses geographic information system (GIS) based data to model the harvest scheduling model. This data partitions the forest into small units for which area, volume and harvest profit information is available. The area of each unit is generally smaller that the maximum clear cut size specified for the Maximum Area Restrictions, so some groups of adjacent units may be harvested together.

We will define the set of Feasible Clusters (Λ) as all groups of adjacent units whose combined area does not exceed the maximum clear cut size. All of these clusters will be generated a priori by enumeration. This can be done efficiently as the maximum area restrictions generally limit the number of units in a cluster to 4 or 5 (Goycoolea et al. 2003). We will say that two clusters are incompatible if they share a unit or if they are adjacent. Forbidding the simultaneous harvesting of incompatible clusters will assure compliance with the maximum clear cut restrictions. This requirement is modeled by Goycoolea et al. (2003) using maximal cliques to impose incompatibilities. These restrictions give the formulation integrality properties that make it relatively easy to solve. Almost all instances of the single period problem are solved to optimality in the root Branch & Bound (B&B) node by CPLEX 8.1.

CLUSTER PACKING HARVEST SCHEDULING MODEL

The Cluster Packing Problem uses geographic information system (GIS) based data to model the harvest scheduling model. This data partitions the forest into small units for which area, volume and harvest profit information is available. The area of each unit is generally smaller that the maximum clear cut size specified for the Maximum Area Restrictions, so some groups of adjacent units may be harvested together.

We will define the set of Feasible Clusters (Λ) as all groups of adjacent units whose combined area does not exceed the maximum clear cut size. All of these clusters will be generated a priori by enumeration. This can be done efficiently as the maximum area restrictions generally limit the number of units in a cluster to 4 or 5 (Goycoolea et al. 2003). We will say that two clusters are incompatible if they share a unit or if they are adjacent. Forbidding the simultaneous harvesting of incompatible clusters will assure compliance with the maximum clear cut restrictions. This requirement is modeled by Goycoolea et al. (2003) using maximal cliques to impose incompatibilities. These restrictions give the formulation integrality properties that make it relatively easy to solve. Almost all instances of the single period problem are solved to optimality in the root Branch & Bound (B&B) node by CPLEX 8.1.

Maximize \( \sum_{S,t} c_{S,t} x_{S,t} \) for each maximal clique \( K \) and for each period \( t \)

subject to \( \sum_{S \in K} x_{S,t} \leq 1 \) for each unit \( u \)

\( \sum_{S \in \Lambda} x_{S,t} \leq 1 \) for each cluster \( S \in \Lambda \) and for each period \( t \)

where:

- \( c_{S,t} \) is net present value of the profit of cluster \( S \) for period \( t \)
- \( \Lambda(K) \) is the set of all clusters that intersect maximal clique \( K \)

This formulation preserves most of the good properties of the single period formulation and is easily solvable, as the computational results will show.

The multi-period model can be complemented with different kinds of restrictions on the volume harvested in each period. The most common restrictions include the production smoothing volume constraints and upper/lower bounds over the volume production.

One typical restriction on the harvested volume is to require total volume in a period to be within \( \pm \Delta \% \) of previous periods. This can be achieved by adding the following restrictions to the multi-period model for each time period \( t > 1 \):

\[
(1 - \frac{\Delta}{100}) \sum_S v_{S,t-1} x_{S,t-1} \leq \sum_S v_{S,t} x_{S,t} \leq (1 + \frac{\Delta}{100}) \sum_S v_{S,t-1} x_{S,t-1}
\]

where \( v_{S,t} \) the volume harvested if cluster \( S \) is selected to be harvested in period \( t \).
Other restrictions that are frequently applied are minimum and maximum harvested volumes. This can be achieved by adding the following restrictions to the multi-period model for each time period $t$:

$$L \leq \sum_{S} v_{S,t} x_{S,t} \leq U$$

where $U$ and $L$ are the maximum and minimum volume allowed to be harvested in each period.

For both types of restrictions it is common that one of the inequalities is active, and hence acts as a fractional generating cut on the LP polytope. This fractional generating effect causes solutions to the LP relaxation to have many fractions. Furthermore, these fractions are difficult to eliminate by variable fixing. As the computational results will show, this makes the problem very difficult to solve.

**ELASTIC VOLUME CONSTRAINT MODEL**

One technique that can be used to minimize the fractional generating effects of volume constraints is to use an elastic version of the constraints. An elastic constraint allows a violation of the restriction, but penalizes this violation in the objective function. In this manner the volume constraints will no longer act as strong cuts, and hence, will generate almost no new fractional extreme points to the LP polytope. This will restore practically all the integrality properties of the multi-period model without volume constraints. Elastic constraints have been successfully used in similar problems (see Ehrgott and Ryan 2003).

It is very difficult to find penalties that will lead to integer solutions that do not violate volume restrictions. For this reason it is a good idea to start penalizing before the restrictions are really violated. So, for example, if we wanted to solve the problem with $\pm 15\%$ production smoothing volume constraints, we could add a $\pm 14\%$ production smoothing volume constraint, allow violations to these constraints, and penalize their violation in the objective function. In this way, if we just keep the violations controlled (below 1%), we will be complying with our target 15% volume constraint.

In the following section we will describe the elastic constraints for the production smoothing volume constraints. The corresponding relaxations for the upper/lower bound volume constraints are analogous.

---

5 i.e. with $\Delta=15\%$ in the original model
6 i.e. with $\Delta_{E}=14\%$ in the elastic model

---

Formulation 2—Multi-period period cluster packing problem with elastic volume constraints

If we add elastic volume constraints to the multi-period model, we obtain the following formulation:

**SHAPE \* MERGEFORMAT**

Formulation 2. Multi-period period cluster packing problem with elastic volume constraints

**INTEGER ALLOCATION**

Although penalties can be easily adjusted to control volume constraint violations for the root B&B node, it might be very difficult to do this and get integer solutions. General purpose LP based heuristics tend to have problems generating solutions with small volume constraint violations. For this reason a custom integer allocation heuristic was developed. The heuristic fixes variables and re-solves the linear relaxation of the model while trying to account for any violations that are too big.

The elastic volume constraints are crucial for the performance of the heuristic. The fractional generating effect of the volume constraints makes it very difficult to develop an LP based heuristic for the strict volume constraint model. Fixing some fractional variables to integrality in this model generally ends in the appearance of an alternate set of fractional variables, making the integer allocation process very slow. This does not happen with the elastic constraint model as the fractional generating effect of the strict volume constraints is not present. On the other hand, if the penalties are big enough, the violations will probably be reasonably controlled. Some corrections of the violations are still necessary, but they are very few due to the penalties.
COMPUTATIONAL RESULTS

Computational tests were run over two instances: a real forest in Northern California called El Dorado and a randomly generated square grid with 144 units. Table 1 shows a summary of the problem characteristics.

Multi-period applications containing 12 and 15 periods were tested for both instances. The runs were made on a Pentium 4 2.0Ghz PC with 2.0 Gb of RAM running Linux. CPLEX 8.1 was used as the MIP solver and problem generation and heuristics were programmed in C++.

Multi-period model without volume constraints

Table 2 shows computational results for the multi-period model without volume constraints. A time limit of 4 hours was imposed, but for all instances it was possible to declare optimality long before that time limit. The first two columns show the instances characteristics. Columns 3 and 4 show the time and B&B nodes needed to declare optimality. Finally the last three columns show information regarding integer feasible solutions found before declaring optimality. Column 5 shows when the first solution with an LP gap \( \leq 1\% \) was found and columns 6 and 7 show the time the first feasible solution was found and its LP gap.

The integrality properties of this model help CPLEX 8.1 find feasible solutions very quickly and also declare optimality in little time.

Production smoothing volume constraint model

Table 3 shows the results for the production smoothing volume constraint model as solved directly by CPLEX 8.1. All tests for this table were run for 8 hours. The format of table 3 is similar to that of table 2. Additionally column 3 shows the level used for the volume constraints. As optimality could not be declared, columns 5 and 6 show the time the best feasible solution was found and its LP gap. Finally column 4 shows the total number of B&B nodes processed in the allotted time. A dash (-) indicates that a feasible solution with the required characteristics was not found.

It can be seen that CPLEX has a lot of trouble finding integer solutions. Although eventually it does find good solutions for El Dorado, computational effort is significant. No integer solutions are found for the grid instances.

Production smoothing elastic volume constraint method

Table 4 shows the results for the elastic constraint method. This method is essentially B&B over the multiple penalties elastic constraint model with constraint branching plus the integer allocation heuristic.

\[ \text{gap} = \frac{(\text{obj}_{\text{best}}_{\text{lp}} - \text{obj}_{\text{ip}})}{\text{obj}_{\text{ip}}} \times 100, \text{where} \text{obj}_{\text{best}}_{\text{lp}} \text{is the greatest linear relaxation optimal value among the B&B nodes to be processed.} \]
\[ \text{obj}_{\text{ip}} \text{is the objective value of the particular integer feasible solution.} \]
The format of table 4 is the same as table 3 with the exception of the meaning of ∆ and how the gaps are calculated. ∆ corresponds to the strict volume constraint we are trying to comply with. Again we use ∆E=(∆-1)% and allow only 1% violation to solve the exact volume constraint with level ∆%. LP gaps are calculated with respect to the LP solution of the corresponding exact ∆% volume constraint model, so they can be compared to the gaps reported in table 3.

Penalties for each constraint are set independently so that the root LP has less than 1% violation, but they are then kept fixed in the B&B tree. A time limit of only 4 hours, instead of 8, was used for these tests.

Although with this method fewer B&B nodes are processed, we can get good solutions quickly for El Dorado and we can also get integer solutions for the grid instances quickly. However, their quality is not good. It should be noted though that the grids where purposefully generated so that it is very difficult to get integer solutions that comply with the volume constraints tightly. Thus, large gaps between the IP and LP solutions for the grid cases are expected. If we compare these results with table 3, we see that the elastic constraint method is much faster than the strict volume constraint model. Integer feasible solutions with similar objective values are found up to 150 times faster\(^8\) with this method.

**CONCLUSIONS**

By eliminating the fractional generating effect of the strict volume constraints, it is much easier to obtain integer feasible solutions. For this reason the elastic constraint method allows good solutions to be obtained much earlier.

---

\(^8\) CPLEX was run for 24 hours for the strict volume constraint model for the ran12by12 instance with 15 periods. Only one solution with a 9% gap was found after 22 hours.
than when solving the strict volume constraint model directly.

It should be noted also that restrictions on harvested volume are generally guides instead of strict requirements, so small violations would likely be acceptable. It is clear that allowing these small violations (for example by allowing violations slightly over 1% of the 14% volume constraint) will give superior results. This provides yet another reason for not using strict volume constraints.

During the computational analysis, it was found that the integer allocation heuristic worked better when the initial LP had little or no violations of the target volume constraints. Because of this, it might be useful to adjust penalties each time a volume restriction is violated in the B&B tree. This would also guarantee that integer solutions found in leafs of the B&B tree would comply with the target volume constraints. We are currently implementing this dynamic adjustment of penalties to be added to the B&B based integer allocation method.

ACKNOWLEDGEMENTS

Partial funding for this research was received from FONDECYT under grant number 1000959. Research support for the second author was provided by the National Science Foundation (Geography and Regional Science Program and the Decision, Risk, and Management Science Program) under grant BCS-0114362.

LITERATURE CITED


EXTREME POLICIES MODELED WITHIN THE LANDSCAPE MANAGEMENT POLICY SIMULATOR (LAMPS)

Pete Bettinger\textsuperscript{1} and Marie Lennette\textsuperscript{2}

\textbf{ABSTRACT}

Several variations on the current behavior of four major landowner groups in the Coast Range of Oregon were simulated using the LAMPS model. The simulation of current and future behavior is termed the Base Case, and assumptions regarding this behavior were derived from numerous meetings with landowner groups associated with the management of Coast Range forests. The extreme policies we model are deviations from the Base Case: limit the maximum clearcut size to 40 acres; set a minimum harvest age of 80 years; assume that entire Coast Range forests are managed by a single landowner group. Results show that minor reductions in harvest levels and net present value are projected when the 40-acre maximum clearcut size is assumed. When the 80-year minimum harvest age is assumed, major reductions in both harvest levels and net present value are projected from Base Case levels. Significant increases are projected for both harvest levels and net present value when we assume that the entire Coast Range is managed by either industrial or non-industrial landowners. These results may follow intuition, but until now have not been quantified for such a large area and long time frame.

\textbf{INTRODUCTION}

The LAndscape Management Policy Simulator (LAMPS) was developed within the CLAMS project (CLAMS 2003) to evaluate alternative forest management policies within the Coast Range of Oregon. The Coast Range analysis area of CLAMS contains about 2.8 million ha of land, spanning the area from the Columbia River south to the northern edge of the Siskiyou National Forest, and from the Pacific Ocean east to the Willamette River. The area contains a patchwork of land ownerships, most notably the Siuslaw National Forest, a significant portion of the of the Bureau of Land Management forests in Oregon, the Tillamook State Forest, several large industrial tree farms, and 400,000 ha of small, non-industrial private forestland.

LAMPS was initially designed to enable the simulation of the “Base Case” forest management strategy of four major landowner groups: federal, state, industry, and non-industrial private. Over the past five years, 75-100 meetings with industrial, federal and state stakeholders were held to determine their current and future management intentions and to assess whether the LAMPS simulation process was adequately modeling their behavior. In addition, surveys of industrial and non-industrial management behavior, conducted by the Oregon Department of Forestry, provided valuable information regarding the behavior of these ownership groups.

In addition to modeling the Base Case, much of the CLAMS modeling work over the past five years has been devoted to modeling minor variations to these policies. This work has been guided by the Oregon Department of Forestry and the Oregon Board of Forestry. Emphasis has been placed on understanding the impacts of potential changes to policies, to allow both managers and policy makers to think through the policies prior to making...
METHODS

LAMPS is a simulation model that allows one to simulate separately the policies of the four major landowner groups in the Coast Range of Oregon. Details of the processes and opportunities for devising alternative management policies in LAMPS can be found in Bettinger and Lennette (2004). Details regarding the mathematical structure of the LAMPS simulation processes can be found in Bettinger and others (2005). We next briefly describe the spatial database structure required for LAMPS simulations as well as a brief description of the scheduling processes for federal, state, industrial, and non-industrial management.

The level of spatial detail required for a scheduling process such as LAMPS is generally negotiated among planning teams. Within the CLAMS project, it was deemed important to maintain fine spatial detail to facilitate modeling of wildlife habitat and geomorphological processes. Therefore, the team decided to recognize aggregations of pixels that had the same vegetation structure, distance from the stream system, and land allocation. These basic simulation units averaged approximately 2 pixels in size. The number of original pixels available from a raster GIS vegetation database developed using a gradient nearest neighbor approach to classification (Ohmann and Gregory 2002) was in excess of 45 million. The number of basic simulation units modeled in LAMPS is about 23 million. Associated with each basic simulation unit were a number of forest structural conditions, including timber volume, average tree age, quadratic mean diameter, average log diameter, and vegetation class. Management units were created by combining watersheds (developed using a 10 m digital elevation model) with land ownership boundaries and aggregated vegetation polygons (large areas of similar vegetation), and subsequent parcelization of the landscape based on the stream system and ridge lines. This process resulted in the development of approximately 441,000 management units. On average, each management unit contains about 50 basic simulation units.

Management units, containing land managed by a single landowner, can be aggregated up into either clearcut harvest blocks or interior habitat areas using a process based on the area restriction model presented by Murray (1999). The area restriction model is a concept related to the spatial aggregation of management units for spatial forest planning processes. Here, any combination of management units that are considered adjacent for planning purposes (sharing a point, line, or within some proximity of each other) can be combined for simultaneous treatment as long as the combined size does not exceed some threshold. Green-up periods, the time it takes regenerated trees in clearcut areas to reach a certain size, are used in conjunction with spatial scheduling rules to control the placement of activities across a landscape. For example, while we may control the maximum size of clearcuts with an area restriction model, we may also control the placement of subsequent clearcuts by preventing their placement next to previous clearcut until some time has passed (the length of the green-up requirement). Area restriction models have thus been used extensively to control the maximum clearcut size in tactical forest planning. They have also been used to build and maintain habitat for which habitat models suggest need be of a certain size (Bettinger et al. 2002).

Management units may also contain multiple land allocations associated with a single landowner. For example, some of the state management land allocations are based a distance from the stream system. In the case of state management, a single management unit may contain three or more land allocations. The level of forest management allowed is assigned at the land allocation level. For example, one land allocation may allow both clearcutting and thinning, partial cutting within riparian areas, and minimal residual legacy trees in regenerated stands. Another land allocation may only allow thinnings to occur, and no activity in riparian areas. The potential timber harvest volume (and hence net revenue) is assessed by determining the level of allowable activity for each basic simulation unit (based on the basic simulation unit's land allocation), and summed to the management unit level for scheduling of activities.
At a higher level of spatial aggregation, LAMPS recognizes land ownerships (federal, state, industrial, and non-industrial), each of which is simulated separately. And finally, “megashed,” ranging in size up to about 0.65 million ha, are recognized. Given the amount of data tracked at the basic simulation unit level (timber volumes, land allocation, and others, and the status of each land allocation in each time period) and the type of computer available (one with 2 Gb RAM), this disaggregation of the Coast Range into reasonably sized megashed was necessary. Results are then generated for each megashed, then aggregated to describe the impact of policies for the entire Coast Range.

LAMPS modeling processes

LAMPS utilizes a different modeling process for simulating the behavior of each landowner group. After attempts to understand the goals and objectives of each landowner group when viewed in aggregate (e.g., all of the industrial landowners in the Coast Range viewed as a single group), a modeling process was chosen to best represent those goals and objectives. For example, on federal land, under current policy, it is unclear whether an objective exists. A number of constraints were identified, such as (1) only a certain percentage of matrix land could be clearcut each year, (2) each watershed needed to contain a minimum percentage of older forest prior to scheduling clearcuts within that watershed, and (3) clearcuts should be relatively small. Therefore, we use a Monte Carlo simulation to spatially simulate forest management activities over time on federal land, subject to the constraints. We also use a unit restriction model to control adjacency, as described in Murray (1999). State land management seeks to achieve the highest even-flow timber harvest volume over time, subject to several ecological constraints (related to achieving forest structural conditions, and maintaining a distribution of interior habitat areas). LAMPS uses binary search to simulate this behavior, and unit restriction adjacency to control clearcut sizes.

Industrial behavior is also modeled using binary search. Here, we noted that over the last 20-30 years, industrial landowners (as a whole) have tended to harvest a relatively even amount of timber each year, even though individual companies may be seeking to maximize other economic goals. In the industrial management simulation model, management units are blocked to create clearcuts that seek to fit a distribution of clearcut sizes using a dynamic deterministic process (Bettinger and Johnson 2003), which uses the area restriction model described in Murray (1999). The non-industrial simulation process also uses this blocking approach to develop clearcuts of certain sizes, yet schedules activities using Monte Carlo simulation. The best we can gather from the behavior of non-industrial landowners is that their tendency to harvest is either based on timber prices (difficult to project a long way into the future) or landowner age (impossible to determine). The Oregon Department of Forestry developed some relationships that show the probability of harvest as a function of stand age, and we use these relationships in LAMPS to decide whether or not to harvest a management unit each time period.

A number of other aspects of management behavior are modeled in LAMPS. These were determined as important via our discussions with the landowner groups, and can be considered a brief description of the Base Case policy for the Coast Range (Table 1).

The extreme policies are modeled by changing some of the assumptions contained in the Base Case scenario. For example, to model the 40 acre maximum clearcut size policy, we simply limit all clearcuts in each of the simulation processes to a maximum of 40 acres. Previously, clearcuts were allowed to be as big as 120 acres. To model the 80-year minimum harvest age, all other Base Case policy assumptions were held constant while a minimum harvest age of 80 years was imposed on all ownerships. Previously under the Base Case, the minimum harvest ages ranged from 35-80 years, depending on the land allocation. To model the policies where we assume that the Coast Range is managed by a single landowner, we first specified that all of the land in the Coast Range was contained within one landowner group, then applied the management behavior described in the Base Case for that landowner group to the land. The only exception was that Congressionally reserved lands (wilderness areas) were maintained in federal ownership. Making this change in land ownership was relatively easy for industrial and non-industrial scenarios. The federal scenario was problematic - we could not identify late successional or riparian reserves on areas that (in the Base Case) were identified as industrial, non-industrial, or state land. State management behavior requires identifying land allocations as a function of distance from the stream system, which would require significant GIS work. Therefore, modeling all lands as if under state ownership was not pursued here. Further, in the forest industry scenario, the forest industry management intensities, which are generally higher on forest industry land in the Base Case, were applied to all lands (except those mentioned above that were not given a new ownership status).
Table 1—Major assumptions in the Base Case policy, by landowner group.

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Federal</th>
<th>State</th>
<th>Forest industry</th>
<th>Non-industrial private</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum harvest age (yrs)</td>
<td>50</td>
<td>45</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>Green-up period (yrs)</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Maximum clearcut size (acres)</td>
<td>—,a</td>
<td>—,a</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>Riparian optionb</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Leave tree optionc</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

a Limited to the maximum size of a single management unit
b Riparian options: 1 = No harvest within Oregon Forest Practices Act buffers, no harvest of hardwoods within 100 feet of a stream; 2 = allow partial harvest within Oregon Forest Practices Act buffers, yet no harvest of hardwoods within 100 feet of a stream; 3 = allow partial harvest within Oregon Forest Practices Act buffers.
c Leave tree options: 1 = leave two trees consistent with the Oregon Forest Practices Act; 2 = leave 5 or 14 trees per acre per Oregon state lands forest plans.
RESULTS

Timber harvest volumes for the Base Case were projected to be around 2 billion board feet per year for the next 100 years (fig. 1), although only two of the landowner groups simulated had even-flow goals (forest industry and state). The net present value of the Base Case policies for the Coast Range is projected to be approximately $12.765 billion. This takes into account harvest revenue, logging costs, site preparation costs, reforestation costs, and weed control and fertilization costs (where appropriate), and uses a 6% discount rate for each landowner group. When clearcut sizes are limited to a maximum of 40 acres, the harvest levels dropped slightly more than 5% (fig. 2), and net present value declined about 7%, to $11.816 billion. One of the reasons that the maximum clearcut size did not have much of an effect is that the average clearcut size in the Base Case was about 40 acres. Increasing the minimum harvest age to 80 years had a more significant effect on the Base Case (fig. 3), since much of the forest in this area of the Coast Range is significantly less than 80 years of age. The even-flow objective of the industrial land, given the harvest constraints in the first few time periods (due to the increased minimum harvest age), significantly constrained projected industrial harvest volumes. While timber harvest levels fell, on average, about 73% from the Base Case harvest levels, net present value fell almost 86%, to $1.846 billion, due to the low harvest levels in the early time periods.

When the entire Coast Range was assumed to be under the management of a single landowner, some interesting results were noted (fig. 4). First, when simulated as being managed under an industrial management regime, projected harvest levels were significantly higher than the Base Case,
as the older forests on federal and state land now facilitate higher near-term harvest levels, allowing for a high even-flow harvest level. Further, potential harvests on formerly state and federal lands are not as constrained by ecological goals as they were in the Base Case. Harvest levels were projected to be almost double the Base Case, and the net present value of the industrial ownership scenario was projected to be about 119% higher than that of the Base Case. The net present value of the non-industrial ownership scenario was projected to be about 74% higher than the Base Case, and harvest levels, while higher than the Base Case, fall from initial relatively high levels, then increase again in later time periods. We believe this to be a function of the probability of harvest process used in the non-industrial case, which is a function of the average age of the timber in each management unit. Here again, potential harvests on formerly state and federal lands are not as constrained by ecological goals as they were under the Base Case.

When the entire Coast Range was assumed to be managed under federal ownership, we find that projected harvest levels initially decline (from Base Case levels), then increase significantly in later time periods. The federal management scenario is not constrained by an even-flow goal, as are the forest industry and state management policies. The main constraint related to harvesting is that more 15% of a watershed needs to be in “older” forest before any clearcut harvesting can occur. Once this happens (after about time period 4), clear-cut harvests are only constrained by the 15% older forest goal, a limit on the total amount of clearcutting per time period (1%), and unit restriction adjacency constraints, thus projected harvest levels are very high in the later time periods, at times higher than any other scenario we modeled (fig. 4). In addition, all “federal” lands that were not previously in federal ownership were modeled as matrix land allocations, so the true federal restrictions (related late successional reserves and riparian reserves) may have been underestimated here. The projected net present value, in fact, of the federal management scenario, is about 7% higher than the Base Case. Figure 5 shows a composite of all of the extreme policies modeled with the LAMPS simulation model.

DISCUSSION

LAMPS is a simulation model designed to assist managers and policy makers in thinking through potential forest landscape policies prior to implementation. It uses a hierarchical structure to model large-scale, long-term policies, and does so for all landowners contained in a landscape. The modeling framework is, of course, a simplification of reality. However, we have conducted numerous meetings with landowner groups who manage land in the Coast Range to determine the most appropriate course of action for modeling their behavior. Although validating such a complex simulation model is problematic, modeling current and future management behavior as close as possible to the actual behavior lends credibility to the results.

One of the major concerns of the LAMPS modeling process is the use of an even-flow goal on state and industry land. The even-flow goal significantly constrains harvest levels in some of the scenarios modeled. Standing timber volumes, in fact, generally increase over time on lands simulated with this goal. Higher total timber harvest volumes may be simulated if the even-flow goal was relaxed. Most of the simulations show a “bottleneck” period that constrains higher even-flow harvest levels. We are currently developing and testing processes to allow upward

Figure 5—Projected timber harvest levels for all forest landscape policies for the Coast Range of Oregon.
deviations in even-flow harvest levels, leaving the con-straining time period at perhaps lower harvest levels. These variable harvest levels will first ensure that the maximum even-flow volume can be achieved, then allow additional harvest without sacrificing volume from any even-flow constraining time period.

Neither the even-flow assumption nor the constraints modeled here as "extreme policies" (40-acre maximum clearcut size or 80-year minimum harvest age) are Law. The even-flow goal was obtained from evidence of recent landowner behavior. Therefore, it seemed to be an appropriate indicator of the behavior of two large landowner groups. Some might argue that in the past, industrial landowners in Oregon had the ability to use federal timber sales to buffer changes in timber markets. It remains to be seen whether this is still possible given the sharp, and recent, decline in federal timber sales. Therefore, the even-flow behavior modeled on state and industrial land may, in the future, change, and give way to a more erratic level of harvest based on maximization of economic or ecological criteria.

These extreme policies that we have modeled with LAMPS provide a perspective on a portion of the solution space that usually goes unexplored in policy analyses. More likely, when developing long-term strategic plans or evaluating the potential effects of new policies, a Base Case is modeled, and minor variations around the Base Case are examined, each reflecting likely changes to regulatory or organizational policy. We feel that by exploring other areas of the solution space, a more complete picture of the productive capacity of the Coast Range forests can be understood.

LITERATURE CITED


This page is intentionally left blank.
A RELATIONAL DATA MODEL FOR
THE GENERATION OF LARGE-SCALE
FOREST SCHEDULING PROBLEMS

Silvana R. Nobre¹ and Luiz C. E. Rodriguez²

ABSTRACT

In Brazil, several pulp and paper companies manage large tracks of land with forest plantations. The management of these forests demands integrated database management systems. One of the functions of these systems is to provide organized information for linear programming harvest scheduling models. This paper suggests a relational data structure that facilitates the analysis, creation and storage of these models and results. Considering that users are not essentially experts in modeling and linear programming techniques, the relational model has to encapsulate complexities, to automate lots of calculations and to assume several consistency checks. Preferably these database and matrix generators have to provide a friendly environment to generate the model and to analyze the results. Advanced data modeling techniques were used to attend these goals. Basically, we describe: (i) the modeling techniques and required parameterization to represent the forest management regimes; (ii) the concept of “calculation batch” to encapsulate the linear programming routines; and (iii) the approach used to set together groups of parameters to define input data, different analysis scenarios and respective results. The approach presented in this paper develops a tool that integrates data modeling and optimization techniques contributed in Brazil for the adoption of non-trivial forest planning techniques.

KEYWORDS: Relational database, harvest scheduling, linear programming, information technology.

INTRODUCTION

The optimization techniques used in forest scheduling models are highly dependent on information technology. Brazilian foresters responsible for the forest scheduling of large pulp and paper companies have had difficulties dealing with database manipulation since they are not essentially experts in modeling and in mathematical programming techniques. Therefore, they often need IT support for making possible the use of mathematical optimization techniques.

These professionals often have to create flexible scenarios and usually demand efficient decision support tools. These demanded tools must encapsulate the complexities of large-scale forest scheduling problems, facilitating the analysis, the creation and the storage of the parameters and the results produced by the several different scenarios studied.

Forest scheduling software ideally should be perfectly integrated to a relationally structured database. And integration, in such cases, means data, parameters and results directly read from and stored in the database. It is also desirable to make all results available to further analysis and interpretation along other information managed by other corporate systems.

The modeling data techniques suggested in this paper are intended to rationally store parameters and all data necessary to generate linear programming type I model matrices, as well as the results produced by the optimization.
process. The techniques, which make the system flexible and easy-to-use, are introduced in the following sections.

**MODELING A FOREST SCHEDULING DATABASE – BASIC CONCEPTS**

This paper considers that the forest scheduling problem can be modeled according to the model I type formulation presented by Johnson and Scheurman (1977). Its basic formulation can be stated as in Clutter and others (1983), considering the following quantities:

- \( N \) = number of forest units
- \( M \) = number of management regimes
- \( T \) = number of management periods in the planning horizon
- \( A_i \) = area of forest unit \( i \)
- \( X_{ik} \) = area of forest unit \( i \) assigned to management regime \( k \)
- \( D_{ik} \) = value (per unit of area) of management regime \( k \) in forest unit \( i \)
- \( V_{iktp} \) = volume (per unit of area) of product \( p \) harvested from forest unit \( i \) in management period \( t \) if management regime \( k \) is used
- \( V_{Min}p \) and \( V_{Max}p \) = minimum and maximum volumes of product \( p \) in period \( t \)

Then, the problem becomes:

\[
\text{Maximize } Z = \sum_{i=1}^{N} \sum_{k=1}^{M} D_{ik} X_{ik} \tag{1}
\]

subject to:

\[
\sum_{k=1}^{M} \sum_{i=1}^{N} X_{ik} \leq A_i \quad (i = 1, 2, \ldots, N) \tag{2}
\]

\[
\sum_{i=1}^{N} \sum_{k=1}^{M} V_{iktp} X_{ik} \geq V_{Min}p \quad (t = 1, 2, \ldots, T) \; (p = 1, 2, \ldots, P) \tag{3}
\]

\[
\sum_{i=1}^{N} \sum_{k=1}^{M} V_{iktp} X_{ik} \leq V_{Max}p \quad (t = 1, 2, \ldots, T) \; (p = 1, 2, \ldots, P) \tag{4}
\]

The presentation of a relational database, flexible enough to generate forest models based on equations (1) to (4), is the main objective of this paper. Although used as the basis for the generation of the coefficients in the above model, a few adjustments in this relational model will easily allow for the consideration of more constraints, and even for the generation of model II type formulations.

Powerful database techniques were used to generate the relational database models described in section 3. Four concepts have to be initially introduced for a better understanding of these techniques: (i) scenario-setting parameterization; (ii) event-output decision-tree; (iii) calculation batch; and (iv) sequencing of activities in a single database.

**Scenario-setting parameterization**

Seventeen different parameters are defined to support the generation of a forest scheduling model. Some parameters are numbers, such as interest rates, planning horizons, first year in the planning horizons etc. Parameters can also define event types, between events sequential rules, output types, and management regimes evaluation methods. Other parameters can yet be tables, such as: management costs, production costs, forest management units, events sequencing rules, production tables etc. The database was designed to make scenarios and parameters all related. That is, a single database where every new scenario and any previous analysis or scenario settings can be recovered altered and restored.

**Event-output decision-tree**

The diverse ways to manage a forest management unit, i.e. the alternative regimes, can be expressed in a decision-tree. The decision-tree is graphically represented through nodes and arcs. Nodes are the relevant events, which alter the forest production structure. Arcs can indicate what happens in between two events.

The initial node of the decision-tree is the last event occurred in the forest management unit, from which other events can occur and so forth. The decision-tree represents what can be done in a forest management unit along a given planning horizon, begins a few years before the beginning of the planning horizon and it can finish before or after the ending of the planning horizon.

The building of the decision-tree is a parameterization process. The parameters, defined by the user, determine the way the events are sequenced, and the length and quantity of possible regimes in the decision-tree. Figure 1 shows an example where two events are defined: clear-cut followed by sprouting (CS) and clear-cut followed by the renewal of the plantation (CR). In the example, the first event occurs three years before the beginning of the planning horizon, when a new rotation of the plantation sprouts from the stumps left by the clear cut.

The decision-tree depicted in Figure 1 was obtained after setting a few parameters in the database which mainly generated: (i) clear cuts followed by the complete renewal of the plantation, i.e., forest cycles of only one rotation; (ii)
coppice regeneration prescriptions, i.e., forest cycles of two rotations where the plantation regenerates from sprouting stumps after a first clear cut; (iii) forest rotations of either 7 or 8 years; and (iv) a planning horizon of 21 years.

Once the decision-tree is defined, the system determines the outcome consequences of the events along each branch. These outputs can be products, revenues, costs, measurable environmental impacts, or even demands. In a forest scheduling type I model, the most common outputs are usually the volumes or weights of certain products.

The sequence of several outputs, one for each node or point along the arcs of the decision-tree, is called output flow. Nodes and arcs are stored in the database, along with all needed data to calculate the output flow, so that each management regime can be completely evaluated. That includes forest age at each node and other essential data demanded by the calculation batch, which offers an environment to define programming sentences and other biometric routines that will produce the real output values.

The correct generation of the event-output decision-tree is crucial, given that the generation of the technical coefficients for the model I type forest scheduling matrix is totally dependent on this concept.

**Calculation batch**

A calculation batch is a sequence of calculation steps. Each calculation batch can be defined in order to satisfy the goals of a given scenario analysis. The batch is composed by a logical sequence of steps. There are four different step types: (i) steps for the creation of temporary tables, (ii) steps for processing SQL statements, (iii) steps to define variables, and (iv) steps to define functions.

The steps that create tables simulate an environment where table subsets can be created from existing tables in the database. The user can further manipulate these subset tables on the following steps of the calculation batch. The SQL processing steps execute SQL sentences on any available table. And the variables and functions definition steps calculate output levels.

Throughout these steps, there is always a target calculation output to which a formula is assigned. Simple formulas can be used to calculate simple output quantities along the output flow. For example, in the statement “IF PW = 1 THEN RESULT:= EST_VOL”, the system checks if “PW” and “EST_VOL” are variables defined in other tables to finally assign the desired result. Complex formulas, involving several variables, internal functions and operators can also be defined, like the one below:

```plaintext
IF PW = 1 THEN BEGIN
    AGE1:= AGEINV * 12 ;
    AGE2:= AGE * 12 ;
    X1:= LN(G) * (AGE1 / AGE2) ;
    X2:= 1 - (AGE1 / AGE2) ;
    X3:= X2 * S ;
    B2EST:= EXP(X1 + B0 * X2 + B1 * X3) ;
    RESULT:= EXP(B2 - B3 / AGE2 + B4 * S + B5 * LN(B2EST)) ;
END
ELSE RESULT:=0;
```

Figure 1—A decision-tree representative of possible regimes considering two relevant regimes: clear-cut followed by sprouting (CS) and clear-cut followed by the renewal of the plantation (CR).
Function definition steps are used to reference available system functions. The following example shows a call to the function that generates an event-output decision-tree.

[Function]
Generate Decision-Tree

Function definition steps might also need parameters, such as the one expressed below that calculates net present values (NPV). In this case, the parameter is a formula that indicates which and how output variables, in the output flow, are passed to the NPV function (where \( V_{PW} \), \( V_{SW} \) and \( V_{BL} \) are values for pulpwood, sawn wood and bare land).

[Function]
Calculate NPV
[Formula]
\[ V_{PW} + V_{SW} + V_{BL} - COST \]

In fact, the calculation batch guides the process of generating the forest scheduling mathematical model. There are required steps that make the process standard, but the possibility of having user defined steps also makes the process very flexible.

Sequencing of activities in a single database
The set of required steps in the calculation batch can be grouped and sequenced as follows:
- Creation of the decision-tree for each forest management unit;
- Valuation of each management regime;
- Generation of the forest scheduling mathematical model coefficients;
- Preparation of the input data for the solver; and
- Conversion of the optimal solution produced by the solver into a forest schedule.

The user can prepare different calculation batches and evaluate several scenarios. The system allows for the creation of different decision-trees, and consequently different forest scheduling mathematical models which will generate different forest schedules.

At each step in the calculation batch, all involved elements are read and stored in the database. In order to give flexibility to the process, each step in the calculation batch must be independent from others. With this approach, which is referred in this paper as the independency principle, it is possible to stop and start the processing of the calculation batch at any step, and also to create intermediate steps without compromising the whole process (Figure 2).

The independency principle imposes the storage of intermediate results in the database, which are: generated decision-trees, output flows, information produced by the solver and final forest schedules. The main modeling techniques used to implement all four described concepts are presented in the next section.

MODELING TECHNIQUES

The most accepted and commercially used approach to project a database system is the Relational Model. A conceptual tool called Entity-Relationship Model (De Miguel and others, 2000; Korth and Silberschatz, 1995; Date, 1991; Setzer, 1990; DeMarco, 1989) is used in this paper to represent relational database structures. In this paper, simple E-R models are used to represent the logical relationships among parameters, intermediate results and final results.

Scenarios and Parameters
Each one of the seventeen types of parameters needed to generate the forest scheduling model has its own structure. All parameters have to be related to the Scenarios entity. A modeling technique called generalization was used to deal with this situation.

A generalization is used to emphasize similarities among entities and to hide their differences. The entity grouping the similarities is defined at a higher level, and the entities representing differences are defined at a lower level. The higher level entity holds the common attributes, and the lower level entities hold the specific attributes. Then the generalization becomes the relationship between the higher level entity and lower level entities (Setzer, 1990).

An entity named ParameterSet was created at a higher level to relate parameters and scenarios. The lower entities are the seventeen different types of parameters. Actually, when this model is implemented, the entity ParameterSet
becomes a table holding all parameters set by the user. Each record in the table refers to one single parameter. Each lower entity becomes one or more tables holding the rest of the attributes of this single parameter. For example, a table of management costs is a parameter with its identification code, name and general description stored in a higher entity. The actual management costs are stored in a lower entity with a much more adequate structure to hold cost tables.

For a given analysis, one scenario refers to a single set of different parameters. And one single parameter, once set, can be used by different scenarios. This is known as a many-to-many relationship between the entity ParameterSet and Scenario. It demands one intermediate table called Scenario_ParmSet in which the relationships between scenario and parameter are registered.

The user can create several different values for each parameter type, and later associate each parameter value to a specific scenario. Figure 3 shows the E-R generalization technique applied to relate the ParameterSet and Scenario entities.

**Event-Output Decision-Trees and the Output Flows**

Each node must be related to the previous node to adequately represent the event-output decision-tree. The only node that does not have an associated previous node is the initial node. A database technique called *auto-relationship* was used to model the association of an entity with itself and to represent the relationship among nodes in the event-output decision-tree.

One set of nodes forming an event-output decision-tree must be related to one scenario and one management unit. In fact, each node ends up related to one single management unit and one scenario.

The final node in each branch of an event-output decision-tree, i.e. the one finishing a sequence of events, relates specifically to the entity *Regime*. It is not necessary to relate
this entity with all nodes representing it. Knowing the last node is sufficient to identify all others because of the auto-relationship technique utilized to represent the output flow. Figure 4 depicts the E-R model used to design the event-output decision-tree logical structure.

With the structure presented in Figure 4, it is possible to easily build the output flow of any management regime predicted for any management unit in a given scenario. As an example, let’s consider only one of the generated regimes, for a certain forest management unit, presented in Figure 1, and now depicted in Figure 5. Let’s also define the following outputs: (i) production level of pulpwood L_PW; (ii) total cost value V_TC; (iii) 7-year old wood stock level L_WS; and (iv) value of the fixed carbon on wood V_FC.

The regime’s output flow illustrated in Figure 5 would generate the records reported in Table 1. Output flows are created to each scenario, and column values are calculated according to the parameters and functions defined in the calculation batch.

Table 1—Records related to the output flow generated by the management regime shown in Figure 5.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Forest Unit</th>
<th>Arc</th>
<th>Year (years)</th>
<th>L_PW (m$^3$/ha)</th>
<th>V_TC (U$/ha)</th>
<th>L_WS (m$^3$/ha)</th>
<th>V_FC (U$/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01 01</td>
<td>0 – 1</td>
<td>1</td>
<td>4</td>
<td>120</td>
<td>203</td>
<td>210</td>
<td>420</td>
</tr>
<tr>
<td>01 01</td>
<td>0 – 1</td>
<td>2</td>
<td>5</td>
<td>147</td>
<td>293</td>
<td>210</td>
<td>511</td>
</tr>
<tr>
<td>01 01</td>
<td>0 – 1</td>
<td>3</td>
<td>6</td>
<td>177</td>
<td>21</td>
<td>210</td>
<td>616</td>
</tr>
<tr>
<td>01 01</td>
<td>0 – 1</td>
<td>4</td>
<td>7</td>
<td>210</td>
<td>21</td>
<td>210</td>
<td>735</td>
</tr>
<tr>
<td>01 01</td>
<td>0 – 1</td>
<td>5</td>
<td>8</td>
<td>240</td>
<td>2030</td>
<td>215</td>
<td>840</td>
</tr>
<tr>
<td>01 01</td>
<td>1 – 2</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1080</td>
<td>215</td>
<td>0</td>
</tr>
<tr>
<td>01 01</td>
<td>1 – 2</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>388</td>
<td>215</td>
<td>0</td>
</tr>
<tr>
<td>01 01</td>
<td>1 – 2</td>
<td>8</td>
<td>3</td>
<td>90</td>
<td>208</td>
<td>215</td>
<td>315</td>
</tr>
<tr>
<td>01 01</td>
<td>1 – 2</td>
<td>9</td>
<td>4</td>
<td>121</td>
<td>203</td>
<td>215</td>
<td>427</td>
</tr>
<tr>
<td>01 01</td>
<td>1 – 2</td>
<td>10</td>
<td>5</td>
<td>150</td>
<td>203</td>
<td>215</td>
<td>525</td>
</tr>
<tr>
<td>01 01</td>
<td>1 – 2</td>
<td>11</td>
<td>6</td>
<td>180</td>
<td>21</td>
<td>215</td>
<td>630</td>
</tr>
<tr>
<td>01 01</td>
<td>1 – 2</td>
<td>12</td>
<td>7</td>
<td>215</td>
<td>1050</td>
<td>215</td>
<td>749</td>
</tr>
<tr>
<td>01 01</td>
<td>2 – 3</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>1080</td>
<td>220</td>
<td>0</td>
</tr>
<tr>
<td>01 01</td>
<td>2 – 3</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td>388</td>
<td>220</td>
<td>0</td>
</tr>
<tr>
<td>01 01</td>
<td>2 – 3</td>
<td>15</td>
<td>3</td>
<td>92</td>
<td>208</td>
<td>220</td>
<td>322</td>
</tr>
<tr>
<td>01 01</td>
<td>2 – 3</td>
<td>16</td>
<td>4</td>
<td>123</td>
<td>203</td>
<td>220</td>
<td>434</td>
</tr>
<tr>
<td>01 01</td>
<td>2 – 3</td>
<td>17</td>
<td>5</td>
<td>152</td>
<td>203</td>
<td>220</td>
<td>532</td>
</tr>
<tr>
<td>01 01</td>
<td>2 – 3</td>
<td>18</td>
<td>6</td>
<td>188</td>
<td>21</td>
<td>220</td>
<td>658</td>
</tr>
<tr>
<td>01 01</td>
<td>2 – 3</td>
<td>19</td>
<td>7</td>
<td>220</td>
<td>2030</td>
<td>220</td>
<td>770</td>
</tr>
</tbody>
</table>
When completely calculated, the output flow is the basis for the generation of the forest scheduling matrix of output coefficients. It is also the key element to calculate each regime’s value and objective function coefficients. Parameters are defined by the user to inform the system which outputs will be used to evaluate the regimes. For example:

Regime’s Value: \( V_{\text{PW}} + V_{\text{BL}} - V_{\text{TC}} \)

The user can also inform what outputs will be used as constraints in the linear programming model.

Results of the optimization

When the matrix is calculated and generated, it is sent to the Solver. The results are saved by variable type in the database. The decision variables, that is, chosen regimes are saved on the regimes’ table. Actually, the chosen area is saved on the regimes’ table.

As the system creates accounting variables to each one of the constraints, these values are saved on correspondent tables. The events’ annual quantities are saved on the result tables (event type); the outputs annual quantities are saved on the result tables (outputs type).

The system must interpret the Solver’s results; in other words, it needs to determine the events that will happen in each forest management unit from the regime chosen by the Solver. The results of this interpretation will also be saved on tables, as figure 6 illustrates.

Calculation batch

The calculation batch is also data, and as an entity it will be treated in the database. At the moment of the calculation, a calculation batch is applied to a scenario. The batch is entirely independent of the scenario. The user can build as many calculation batches as necessary and apply them to any scenario.

All steps in a calculation batch, required and user defined, and of any type, are saved on a table. Making a calculation batch is simply a matter of choosing instructions from this table and ordering them in a correct execution sequence.

Due to the independence principle, it is possible to create batches that execute only parts of the process. For example, it is possible to initially create a batch that builds the linear programming matrix from the calculated output flow, and to send the matrix coefficients to the solver with one set of constraint levels. Altering only the last few steps in the calculation batch makes it possible to analyze several different sets for the constraint levels. The users can then produce several rapid simulations, store the results and generate as many optimizations as necessary.
CONCLUSIONS

This paper presents a method to model, in one single relational database, the four essential tasks generally needed to model any forest scheduling problem. The four tasks are: parameterization of different scenarios, generation of several management regimes in a decision-tree structure, calculation/evaluation of outputs and final report writing. The definition of a well defined relational structure in one single database encapsulated many of the complexities, and turned possible the development in Brazil, of a large scale forest decision support system (DSS) based on mathematical scheduling optimization techniques.

The system is being used by three large Brazilian pulp mills to manage their forests. Usually, before the acquisition of the DSS built after principles presented in this paper, weeks were needed to develop the complete formulation of the LP problem. Nowadays, formulations with hundreds of thousands of variables are generated, solved and reported in short periods of time varying from two to four hours. A normal solution for problems with approximately four hundred forest units, and one hundred alternative management regimes per forest unit, takes approximately two hours, depending on the computer memory and processor.

Many of the users of forest management decision support systems in Brazil do not have any formal training in mathematical programming techniques, and even more rarely in database management. The approach used to develop the database structure not only turned possible to hide several unnecessarily difficult data management tasks but also generated a consistent environment that helps to avoid information degeneracy, despite the training level of the user.

Rid of undesirable complexities, the user finds the process of using the system less challenging and more educational. The process ends up offering an environment to check for intermediate results, to visualize the output flows and to analyze the impacts of altering the level of the several constraints. Basically, the user becomes really capable of analyzing several different scenarios because it becomes much easier and less time consuming.

LITERATURE CITED


SHORT-TERM HARVEST SCHEDULE SENSITIVITY TO FUTURE STUMPAGE PRICE ASSUMPTIONS

Eric S. Cox

ABSTRACT

Forest planning models have long been used as an analytical tool for providing information to facilitate effective decision making and planning. Inherent to the financial analyses conducted with these models are assumptions concerning key financial parameters contained in the model such as discount rates, future costs, and future stumpage prices. While projecting timber prices into the future with any accuracy is an extremely difficult exercise, price forecasting is nonetheless a critical part of forest planning analyses. The ramifications of these assumptions over a long planning horizon can be significantly different product flows, activity levels, and cash flows. The purpose of this study is to investigate the impact of different future stumpage price assumptions on the short-term (5-year) timber harvest schedule for a southern pine forest, and to examine how much of the schedule is financially driven. The findings indicate that the short-term harvest schedule is sensitive to different price projections. This result is significant especially with respect to the timing of short-term timber harvest decisions to take advantage of market prices.

KEYWORDS: Stumpage prices, harvest schedule, forest planning.

INTRODUCTION

While projecting timber prices into the future with any accuracy is an extremely difficult exercise, price forecasting is nonetheless a critical part of forest planning analyses. For example, it is well known that timber price fluctuations are a significant factor with regard to timberland returns. The ramifications of these assumptions over a long planning horizon can be significantly different product flows, activity levels, and cash flows. The purpose of this study is to investigate the impact of different future stumpage price assumptions on the short-term (5-year) timber harvest schedule. Various stumpage price projections were devised, with the resulting short-term harvest schedules compared for purposes of examining how much of the schedule is financially driven. These price projections were applied to a case study of a southern pine forest to evaluate their influence on short-term timber harvest decisions.
BACKGROUND

The Forest—The (hypothetical) forest modeled for this study is 100,000 acres in size, and consists entirely of loblolly pine (*Pinus taeda*) plantations. There are 126 stands, and the age of these stands ranges from 1 to 30 years. A uniform age class distribution was modeled.

The Stumpage Price Projections—Three future stumpage price forecasts were modeled in this study:
1. Flat real prices over a 100-year planning horizon (*flat prices*).
2. A 1% real annual increase (over and above inflation) in all products over the planning horizon (*increasing prices*).
3. A 1% real annual increase (over and above inflation) in years 1 to 5 for all products except pine sawtimber. For pine sawtimber, there was an equal annual price decrease in years 1 to 5 such that the resulting stumpage price for pine sawtimber would equal the stumpage price for pine chip ‘n’ saw. Prices were then held flat over the remaining years of the planning horizon (*modified prices*).

The Model—A model II linear programming formulation was used to develop the timber harvest schedule for the forest in this study. The LP-based model consisted of an objective function maximizing net present value (NPV) over a 100 year planning horizon composed of 1-year periods.

In developing the harvest scheduling model, several assumptions were made, including: (1) clear-cut stands are site prepped the year following harvest and planted two years following harvest; (2) all stands that are thinned receive a post-thin fertilization the year following thinning; (3) thinning is optional, there is only one thinning per rotation, and thinning can be scheduled for ages 14-20; (4) minimum rotation age is 20; and (5) the financial analysis is before tax using a real discount rate (net of inflation) of 8%.

Growth & Yield—Growth and yield projections by product were developed using a proprietary Forest Technology Group loblolly pine plantation growth and yield model. Per acre harvest volumes generated by the growth and yield model were used as production coefficients in the harvest scheduling model.

RESULTS

A total of six harvest scheduling model runs were conducted for this study, based on the three alternative stumpage price scenarios and two alternative model formulations: a model constrained to produce a positive cash flow (net revenue) of greater than or equal to $25 million in each of years 1 to 5, and a model without this cash flow constraint (unconstrained). The six model results were used to evaluate the sensitivity of the short-term harvest schedule to different stumpage price projections. Comparison of the results provides valuable insight concerning the extent to which the short-term harvest schedule is financially (price) driven.

Long-Term Results—A brief look at some long-term results is valuable for gaining perspective into the impact of the different price projections on the timing of thinnings and regeneration harvests, and the mix of forest products produced. The average harvest ages over the first 50 years of the planning horizon are shown in Table 1.

As expected, the rotation age is longest under the increasing prices scenario, and shortest under the modified prices scenario. Under modified prices, there are no thinnings scheduled after year 10, as there is no price premium attached to the production of sawtimber.

Average annual pine harvest volumes by product over the first 50 years of the planning horizon are summarized for both the unconstrained and constrained models in table 2.

As expected, the flat and increasing price scenarios result in a mix of products weighted towards the production of sawtimber (PST), while the modified prices scenario results in a product mix heavy to the production of chip ‘n’ saw (PCNS). Further, due to a shorter rotation with no thinnings, the modified prices scenario results in a greater total pine harvest volume. Lastly, comparison of results between the unconstrained and constrained models shows no appreciable difference.

Short-Term Results—Again, for this study the short-term has been defined to be the first five years of the model. The short-term results to be examined here are harvest acres, harvested stands, silvicultural costs, harvest volumes, and net revenue. Acres clear-cut and thinned under the
different price projections are summarized in tables 3 and 4 respectively.

As expected, the modified prices scenario results in the highest total acres clear-cut and the lowest total acres thinned. Also, clear-cut acres are greater in the constrained model for all three price projections. The increasing prices scenario had the lowest total acres clear-cut under both model formulations.

Comparison of results between the unconstrained and constrained models with regard to acres thinned shows slightly fewer acres thinned in the constrained models. Worth noting for the unconstrained model is that the same thinning acreages are chosen under both flat and increasing prices.

The silvicultural costs under the different price projections and model formulations are summarized in table 5.

For the unconstrained model, total silvicultural costs are highest under the modified prices scenario. This reflects the much higher stand establishment costs associated with this price scenario having the highest number of acres clear-cut.

### Table 2—Average annual pine harvest volumes by product under alternative price projections and model formulations for mixes of sawtimber (PST), chip ‘n’ saw (PCNS), and pulpwood (PPWD) production.

<table>
<thead>
<tr>
<th>Model run</th>
<th>Unconstrained</th>
<th>Constrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>PST</td>
<td>306,085</td>
<td>305,901</td>
</tr>
<tr>
<td>PCNS</td>
<td>162,384</td>
<td>162,622</td>
</tr>
<tr>
<td>PPWD</td>
<td>205,224</td>
<td>205,705</td>
</tr>
<tr>
<td>Total</td>
<td>673,693</td>
<td>674,228</td>
</tr>
</tbody>
</table>

### Table 3—Total and annual acres clear-cut under alternative price projections and model formulations for years 1-5.

<table>
<thead>
<tr>
<th>Year</th>
<th>Unconstrained</th>
<th>Constrained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Flat</td>
<td>Increasing</td>
</tr>
<tr>
<td>1</td>
<td>27,159</td>
<td>24,714</td>
</tr>
<tr>
<td>2</td>
<td>3,087</td>
<td>4,310</td>
</tr>
<tr>
<td>3</td>
<td>417</td>
<td>1,639</td>
</tr>
<tr>
<td>4</td>
<td>3,384</td>
<td>859</td>
</tr>
<tr>
<td>5</td>
<td>3,194</td>
<td>4,192</td>
</tr>
<tr>
<td>Total</td>
<td>37,241</td>
<td>35,713</td>
</tr>
</tbody>
</table>

### Table 4—Total and annual acres thinned under alternative price projections and model formulations for years 1-5.

<table>
<thead>
<tr>
<th>Year</th>
<th>Unconstrained</th>
<th>Constrained</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Flat</td>
<td>Increasing</td>
</tr>
<tr>
<td>1</td>
<td>10,429</td>
<td>10,429</td>
</tr>
<tr>
<td>2</td>
<td>5,635</td>
<td>5,635</td>
</tr>
<tr>
<td>3</td>
<td>4,053</td>
<td>4,053</td>
</tr>
<tr>
<td>4</td>
<td>1,667</td>
<td>1,667</td>
</tr>
<tr>
<td>5</td>
<td>3,333</td>
<td>3,333</td>
</tr>
<tr>
<td>Total</td>
<td>25,117</td>
<td>25,117</td>
</tr>
</tbody>
</table>
Table 5—Total and annual silvicultural costs under alternative price projections and model formulations for years 1-5.

<table>
<thead>
<tr>
<th>Year</th>
<th>Flat</th>
<th>Increasing</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>683,334</td>
<td>683,334</td>
<td>783,335</td>
</tr>
<tr>
<td>2</td>
<td>6,486,808</td>
<td>5,997,916</td>
<td>6,106,918</td>
</tr>
<tr>
<td>3</td>
<td>3,960,007</td>
<td>3,947,785</td>
<td>4,608,023</td>
</tr>
<tr>
<td>4</td>
<td>766,215</td>
<td>1,138,995</td>
<td>1,578,340</td>
</tr>
<tr>
<td>5</td>
<td>987,188</td>
<td>610,470</td>
<td>1,746,381</td>
</tr>
<tr>
<td>Total</td>
<td>12,883,553</td>
<td>12,378,500</td>
<td>14,822,998</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Flat</th>
<th>Increasing</th>
<th>Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>683,334</td>
<td>683,334</td>
<td>783,335</td>
</tr>
<tr>
<td>2</td>
<td>3,428,603</td>
<td>3,433,655</td>
<td>2,964,540</td>
</tr>
<tr>
<td>3</td>
<td>3,043,784</td>
<td>2,869,635</td>
<td>3,108,929</td>
</tr>
<tr>
<td>4</td>
<td>2,470,472</td>
<td>2,571,172</td>
<td>2,762,547</td>
</tr>
<tr>
<td>5</td>
<td>2,456,033</td>
<td>2,583,728</td>
<td>2,877,255</td>
</tr>
<tr>
<td>Total</td>
<td>12,082,226</td>
<td>12,096,525</td>
<td>12,496,607</td>
</tr>
</tbody>
</table>

Silvicultural costs in the constrained model are lower in comparison to the unconstrained model for all three price projections. With flat prices, stand establishment costs are lower due to fewer acres clear-cut during years 1 to 4. Primarily, this reflects lower planting and herbaceous weed control costs. In addition, fewer acres thinned results in lower post-thin fertilization costs.

With increasing prices, the higher number of acres clear-cut during years 1 to 4 resulted in increased site prep costs. But this was offset by lower planting and herbaceous costs, and slightly lower post-thin fertilization costs.

With modified prices, stand establishment costs are significantly lower in the constrained model due to fewer acres clear-cut during years 1 to 4 (about 6400 acres less).

Total and annual pine harvest volumes by product under the different price projections are summarized for both the unconstrained and constrained models in table 6. Additionally, these total harvest volumes by product are shown in figure 1a for the unconstrained model, and figure 1b for the constrained model.

For both model formulations, there is less pulpwood (PPWD) harvested under modified prices due to the lower number of acres thinned, while the higher number of acres clear-cut under this price scenario results in a higher PCNS harvest and a slightly higher PST harvest.

For the unconstrained model, total pine harvest volumes range from 4.9 million tons (increasing prices) to 5.2 million tons (modified prices). For the constrained model, total pine harvest volumes range from 5.6 million tons (increasing prices) to 5.8 million tons (both flat and modified prices).

Thus, constraining the model to meet or exceed a minimum cash flow target results in higher harvest volumes for each product (and, as follows, in total), and a slightly narrower difference in total harvest volume between the different price projections. These results are in line with expectations.

Total and annual net revenue under the different price projections is shown in figure 2a for the unconstrained model and figure 2b for the constrained model. Note that net revenue as reported here is not the objective function value, which is NPV.

For the unconstrained model, total net revenue is highest under the modified prices scenario. This follows from this price scenario having the highest harvest volume, particularly concerning PCNS and PST. Net revenue is negative in year three for all pricing scenarios due to 1) a low number of acres clear-cut and a higher number of acres thinned, and 2) the significant number of acres clear-cut in year one are planted and receive herbaceous treatment in year three. Total net revenue ranges from $129 million (both increasing and flat prices) to $140 million (modified prices).

As described previously, the constrained model employed a minimum positive cash flow constraint covering years 1 to 5. Total net revenue is higher in comparison to the unconstrained model for all three price projections. Total net revenue ranges from $153 million (increasing prices) to $157 million (modified prices), with all of the difference in net revenue occurring in year 1. Following from the earlier outcomes regarding harvest volumes, constraining the model results in higher total net revenue, and a narrower difference in total net revenue between the different price projections (from $11 million to $4 million).
CONCLUSIONS

The results of this study indicate that the short-term harvest schedule is sensitive to the different price projections modeled in both the unconstrained and constrained models. This result is significant especially with respect to the timing of short-term timber harvest decisions to take advantage of market prices. Financial objectives may indicate the need for flexibility concerning targeting short-term harvesting decisions in response to market prices. That is, the timing of harvests with regard to the mix of forest products produced is important, especially as it concerns financial goals.

The price sensitivity is related to both the forest examined in this study and the model formulation of the harvest scheduling problem. The uniform age class distribution of this forest allowed flexibility in relation to the stands scheduled for harvest and the timing of these short-term harvest decisions. The model formulation also provided flexibility. Some examples of this flexibility are that thinning is optional, and age 20 stands could be either thinned or clear-cut. Lastly,
Table 6—Total and annual pine harvest volumes by product under alternative price projections and model formulations for years 1-5.

<table>
<thead>
<tr>
<th>Year</th>
<th>PPWD Flat</th>
<th>PPWD Increasing</th>
<th>PPWD Modified</th>
<th>PCNS Flat</th>
<th>PCNS Increasing</th>
<th>PCNS Modified</th>
<th>PST Flat</th>
<th>PST Increasing</th>
<th>PST Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>722,153</td>
<td>654,352</td>
<td>500,924</td>
<td>1,198,969</td>
<td>1,067,491</td>
<td>1,106,572</td>
<td>1,531,317</td>
<td>1,508,183</td>
<td>1,557,361</td>
</tr>
<tr>
<td>2</td>
<td>234,418</td>
<td>268,907</td>
<td>205,360</td>
<td>201,826</td>
<td>278,260</td>
<td>453,921</td>
<td>101,855</td>
<td>118,566</td>
<td>230,744</td>
</tr>
<tr>
<td>3</td>
<td>95,186</td>
<td>127,031</td>
<td>165,140</td>
<td>3,071</td>
<td>68,838</td>
<td>92,748</td>
<td>0</td>
<td>17,622</td>
<td>13,028</td>
</tr>
<tr>
<td>4</td>
<td>150,324</td>
<td>106,672</td>
<td>113,054</td>
<td>196,491</td>
<td>47,419</td>
<td>293,817</td>
<td>2,30,036</td>
<td>36,985</td>
<td>145,776</td>
</tr>
<tr>
<td>5</td>
<td>145,222</td>
<td>156,068</td>
<td>2,877,255</td>
<td>130,792</td>
<td>195,008</td>
<td>2,877,255</td>
<td>67,125</td>
<td>257,554</td>
<td>79,189</td>
</tr>
<tr>
<td>Total</td>
<td>1,347,303</td>
<td>1,313,030</td>
<td>1,104,814</td>
<td>1,731,149</td>
<td>1,657,017</td>
<td>2,105,187</td>
<td>1,930,333</td>
<td>1,938,910</td>
<td>2,026,099</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>PPWD Flat</th>
<th>PPWD Increasing</th>
<th>PPWD Modified</th>
<th>PCNS Flat</th>
<th>PCNS Increasing</th>
<th>PCNS Modified</th>
<th>PST Flat</th>
<th>PST Increasing</th>
<th>PST Modified</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>310,789</td>
<td>320,668</td>
<td>110,804</td>
<td>330,277</td>
<td>345,046</td>
<td>198,816</td>
<td>1,009,552</td>
<td>965,723</td>
<td>1,174,429</td>
</tr>
<tr>
<td>2</td>
<td>310,018</td>
<td>283,265</td>
<td>163,818</td>
<td>381,305</td>
<td>340,721</td>
<td>408,714</td>
<td>353,200</td>
<td>378,586</td>
<td>364,057</td>
</tr>
<tr>
<td>3</td>
<td>239,549</td>
<td>271,101</td>
<td>288,135</td>
<td>420,467</td>
<td>452,737</td>
<td>533,330</td>
<td>324,869</td>
<td>280,922</td>
<td>270,560</td>
</tr>
<tr>
<td>4</td>
<td>200,645</td>
<td>278,631</td>
<td>219,233</td>
<td>452,835</td>
<td>420,761</td>
<td>575,352</td>
<td>292,002</td>
<td>289,778</td>
<td>255,876</td>
</tr>
<tr>
<td>5</td>
<td>388,624</td>
<td>254,948</td>
<td>398,489</td>
<td>505,272</td>
<td>423,879</td>
<td>636,373</td>
<td>235,333</td>
<td>283,670</td>
<td>177,455</td>
</tr>
<tr>
<td>Total</td>
<td>1,449,625</td>
<td>1,408,612</td>
<td>1,180,570</td>
<td>2,090,156</td>
<td>1,983,145</td>
<td>2,352,585</td>
<td>2,214,955</td>
<td>2,198,679</td>
<td>2,242,379</td>
</tr>
</tbody>
</table>
with regard to the constrained model, the cash flow constraint was not so burdensome as to entirely dictate the solution.

Along these lines, there are several factors worthy of investigation in terms of their impact on the sensitivity of the short-term timber harvest schedule to different future stumpage price assumptions. A few of these factors would include:

1. A skewed age class distribution or age class gaps. Clearly, forest age class structure would be a key driver with respect to price sensitivity. A younger forest with limited merchantable stands in the short-term would likely result in a more biologically driven solution. An older forest with many slow-growing stands would likely result in a more biologically driven solution. That is, the sensitivity to the prices modeled could be muted in both these instances where the forest age class structure dictates the solution.

2. Forest policy constraints. Much like the cash flow constraint, other forest policy constraints are likely to reduce price sensitivity.

![Figure 2 Total and annual net revenue under alternative price projections for a) the unconstrained model, and b) the constrained model.](image)
3. Price increases/decreases by product. As many analyses have confirmed, this can have a significant impact with regard to optimal silvicultural prescriptions. Thus, model sensitivity to price could differ with the price projections modeled.

4. Spatial harvest planning. Large contiguous blocks of the same species of very similar age can significantly affect the results of operational harvest planning due to adjacency issues, perhaps reducing the effect of the prices modeled.

Further investigation of these or other factors would make important contributions to the theme of stumpage price sensitivity of short-term forest planning results.
DECISION SUPPORT IN GERMAN FORESTRY
BY USING INDUSTRIAL SIMULATION SOFTWARE

Martin Hemm¹ and Anne-Katrin Bruchner²

INTRODUCTION

Enterprises in the forest sector today have to compete on a global timber market. To fulfill the requirements of their customers and to compete with international timber prices it is necessary to reduce operational and transport costs in the timber harvesting process.

The challenge of forestry today is to maintain a market oriented timber supply and to become a proactive industry instead of a reactive one. Most harvesting operations are planned for silvicultural regimes, i.e. single trees have to be cut in order to provide space for the remaining ones and to improve the quality of the stand. With growing investments into machines, in the future economical aspects are going to be even more important for the positioning of an enterprise on a global timber market. Low operational costs and high technical productivity of the machinery will become increasingly necessary for forest operations. Improvement of the integral logistics management can only be achieved by taking a close look at all elements of the production chain (Warkotsch, Ziesak 1998).

The main task of forest logistics is to manage the material and information flow in all segments of the wood supply chain. In order to be able to respond to the industrial dynamics it is important for the forest enterprises to know precisely their production layout and stand inventory at any time. Planning processes within the operational system that are based on an up-to-date inventory of the raw material could be optimized by using supporting tools like industrial simulation software. Due to the silvicultural approach in Germany approximate information about the spatial location of the selected trees is necessary.

RESEARCH METHODOLOGY AND APPLIED SIMULATION TECHNIQUE

Evaluation of simulation programs for harvest planning

An evaluation of different forestry-based calculating and planning tools from various countries gave an insight into already existing software applications. In a study made by Hemm (2002) seven simulation programs for supporting the planning process of forestry operations were tested. The investigation included software packages from Canada,

¹ Diplom-Forstwirt (Univ.) Martin Hemm is currently working at the Department of Forest Work Science and Applied Computer Science at the Technical University of Munich, Am Hochanger 13, 85354 Freising, Germany.

² Diplom-Forstwirtin (Univ.) Anne-Katrin Bruchner is now working at the Department of Wood Science and Forest Products, Virginia Polytechnic Institute and State University, USA.
Most of these simulation tools are designed for regional silvicultural specifics and local harvesting problems, which makes it difficult to use them in Germany without any necessary modifications. Due to these results, the department decided to create a new tool, which should assist in finding the most efficient machine combination for harvesting and logging processes and could also be used as an e-learning tool for universities and forest schools.

The next step of the research study was to evaluate simulation software that could be adapted to complex systems such as the forestry production chain. Eight different software packages were compared and evaluated in a multiple goal analysis by means of a catalogue of criteria containing several specific forest requirements. In the beginning the requirements of the programs on production layout, manufacturing process and products were tested.

This step has been followed by an intensive investigation of one selected product, AutoMod™, made by Brooks Automation Inc., USA. AutoMod is parameter-driven and requires computer language based programming as well as visual interactive programming by means of a manufacturing simulator. Within this discrete simulation a material handling system can be defined with all its physical components in an editing environment in which the logic is also programmed. A simulation can then be run in a simulation environment creating a detailed 3-D real-time visualization of the system (Banks, Carson, Nelson 1999).

Simulating realities

Simulation software is designed to analyse, plan and control material handling systems. It is useful for analysing large and complex real-world situations that can not be solved by mathematical operations research models. A large number of variables, parameters and also functions can be handled.

An advantage of modeling is that it does not disturb the real system, and therefore it is possible to test various decision-making scenarios without interference. Additionally it will take only a short period of time to simulate a certain system, because it is possible to compress time. Furthermore simulated data is much cheaper to collect than similar data from a real-world system. A computer based simulation model describes the operation of the system and its individual components whose behaviour can be predicted. This feature allows a study of the interaction between individual processes in the system (Bruchner 2002). To simulate the realities occurring in forestry, there are two sequent steps to take: Modeling of the production plant and modeling of the production process (fig. 1).

Modeling the production plant

The production plant represented by the stand was modeled by using a stand simulator called “SILVA”, developed at the Technical University of Munich (Pretzsch, Biber and Dursky 2002). The research focused on a stand of one hectare size located in Southern Germany. Cruise data was taken from systematic samples placed in this certain stand. Then, on the basis of the cruise data the stand was reproduced in SILVA and a tree list was generated by the program, including information about tree number, tree species, dbh and height as well as x- and y-coordinates for every single tree.

This information, listed in a MS Excel table, has been transferred to AutoMod™ and the stand generated by SILVA was reproduced as production plant in the AutoMod™ virtual reality environment. On the basis of this information, the times of the harvester for positioning its aggregate and processing are calculated for every single tree.

But the information, which is necessary to calculate productivity during the simulation was still missing: the number of assortments of each tree and their volumes.

Creating assortments and volumes

To solve that problem a calculation software called “Holzernte”, developed at the FVA Freiburg, Germany, was applied (Hradetzky, Schopfer 2001). Input data for “Holzernte” is a list of trees to cut during a planned harvesting operation. On the basis of this input data the user is able to define the pile he wants to accumulate. Holzernte
then calculates the assortments and their volume for every specified pile. At the end of the calculating process, a MS Access table including the assortments and volumes is generated.

This table was connected with AutoMod™, which was done by means of a Visual Basic link. Provided with the tree data the model of the one hectare big stand was generated in AutoMod™.

Roadway modeling

One feature in AutoMod™ allows to draw a road system (Banks 1999). In the model there are a rectangular extraction lines and one forest road implemented. The roadways in the model have special attributes, which make it possible to define the roads either as extraction line or as forest road. On the roadways there have been placed control points at regular intervals of one meter. The control points are marks for the machines driving on the roadways, where they can stop and start processing. Next to the forest road two depots were created, in which the logged assortments are stored.

MODELING THE PRODUCTION PROCESS

After modeling the production plant, the next step was to simulate the production processes. Therefore one harvesting (harvester) and one primary transportation (forwarder) scenario had been selected. An interface was generated, in which the user has to define machine parameters. These parameters can be varied for every new simulation run.

The input-interface

For both of the machines, harvester and forwarder, as well as for some general parameters a special input-interface was developed (table 1).

In the following section you will find short explanations of the most important parameters, which are the basic requirements for simulating the harvesting and logging processes (table 2).

Simulation process

Each simulation run starts with the compilation of the model. The trees, roadways and machines are generated in a virtual reality (fig. 2).
Table 2—Explanations of the input-parameters

<table>
<thead>
<tr>
<th>parameter</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>assortment list / percentage trees to cut</td>
<td>There are two different options to identify those trees, which should be cut. The first way is about using an assortment list. The trees of this list have already been chosen for cutting during the stand simulation in SILVA. The second way can be taken by the user in case there is no assortment list defined. In AutoMod™ one is able to fill in a percentage value of the trees to be cut. Then, the number of trees, which accord to the given percentage value is selected by random.</td>
</tr>
<tr>
<td>log length</td>
<td>The log length given in meters can be chosen by the user, if there is no assortment list with predefined log length existing.</td>
</tr>
<tr>
<td>MTBF (mean time between failure)</td>
<td>Describes the time between two breakdowns of harvester or forwarder.</td>
</tr>
<tr>
<td>MTTR (mean time until repair)</td>
<td>Describes the length of the breakdowns by the help of a triangular function with minimum, most likely and maximum values.</td>
</tr>
<tr>
<td>preparation time, time loading and unloading per cycle (forwarder)</td>
<td>Also described by a triangular function.</td>
</tr>
</tbody>
</table>

Figure 2—Virtual Reality environment of the model
After having finished its preparation process, the harvester starts driving on the extraction line to the nearest tree to be cut. It stops on the control point, which is the closest to the tree. There it starts to cut and process the tree and takes down the logs next to the extraction line before moving to the next tree.

The forwarder follows the harvester in a predefined distance and starts hauling the assortments. When the forwarder has reached its maximum loading capacity, it returns to a depot, which is placed near the forest road and unloads the assortments.

During this working process a breakdown may occur. In that case the machine stands still at its current position until the repair is finished. If the harvester is down for a longer time, this also may have an impact on the forwarder’s productivity.
The simulation run ends, when all marked trees are cut and processed by the harvester and all logs are unloaded at the depots by the forwarder.

The model was validated by comparing the output with several time study results in terms of productivity and energy consumption. The data needed to develop this model and also to validate it derives from different time studies made in Germany and Austria. The validation runs were based on two kinds of data variation:

1) the stand data was changed while the machine parameters remained the same
2) machine data was changed while only one specific stand was the basis for the model

The validation proved that the model executes the specified logic and performs all functions as expected. Furthermore the whole process of building the model once more showed how the quality of the results depends on the data that has been used to create it.
RESULTS OF THE MODEL

At the end of each run AutoMod™ creates an output window. Three columns show the results for harvester, forwarder and depot calculated during the simulation (table 3). The specific values can be seen at table 3. Concerning the depot, the information is split into stem wood classes H1 to H5 and industrial wood.

The output-interface delivers the results of one run. In AutoMod™ it is possible to do multiple runs with different settings in the input-interface and in this case multiple output-interfaces are produced. That for example allows the user to compare the output of different scenarios with different types of harvesters and get a decision support.

ADVANCED MODEL

After validation of the first model had been finished the next modeling process included an extension of several attributes. Therefore the following features were integrated:

• enlargement of the model => the simulated stand is expanded to 11 hectares
• inclusion of a digital terrain model
• inclusion of a digital soil map
• modeling of an extraction lines system which was measured and mapped (fig. 3)

• simulation of different harvesting and logging systems => new operations include skidder, horse logging and manual felling (fig. 4). This required a new user interface (fig. 5).

The last part of the project will contain a simulation of different tree species and mixed stands as well as some combinations of different harvesting and logging systems. To complete the whole project successfully, it will be necessary to obtain enough data, which allows detailed calculations and satisfying reproduction of realities.

CONCLUSION

Industrial simulation software can be a flexible tool for modeling production processes in the wood supply chain as well as in pulp, paper and timber industry. Various simulation programs are available but only a few of them could be used to reproduce such complex forestry systems. To model different timber harvesting scenarios it is necessary to create a model of a forest enterprise which represents the production area as close to reality as possible and provide actual inventory data about material, machines and stand conditions. The main purpose is to create an instrument for planning and controlling all production processes.

The AutoMod™ harvest planning model can be used as decision support tool:
The results calculated during several simulation runs contain detailed data on economical, ecological and technical aspects, which are useful for a forest enterprise to see the positive and negative effects of different harvesting systems and logging operations on their budget and the environment.

e-learning component:
At universities and forest schools the impact of system changes can be demonstrated to show students the coherences and effects aroused by their decisions.

LITERATURE CITED


MILL SUPPLY AND FOREST PRODUCT MARKETS
This page is intentionally left blank.
ELASTICITIES OF DEMAND FOR INTERNATIONAL FOREST PRODUCT IMPORTS

Extended Abstract

J.A. Turner, and J. Buongiorno

ABSTRACT

Models of the derived demand for forest product imports were estimated for major forest products, with data from 64 countries for 1970 to 1987, and tested with data from 1988 to 1997. The best results were obtained with a dynamic model estimated by the Arellano-Bond method. For most products the demand for imports was price inelastic and income elastic.

INTRODUCTION

Estimates of price and income elasticities of import demand are needed to predict the effects of tariff liberalization and other policies on international forest product markets (Zhu and others 2001). These elasticities are key elements of models like the Global Forest Products Model (GFPM, Buongiorno and others 2003).

The utilization of international panel data (Chou and Buongiorno 1984; Uusivuori and Kuulivainen 2001) have helped generalize import elasticity estimates, but with attendant difficulties, such as bias and inconsistency in dynamic models estimated by fixed or random effects (Hsiao 1986).

The main objective of this study was to better estimate the demand for forest product imports. Panel data from 64 countries for 1970 to 1987, was used to estimate static and dynamic models by five estimation methods. The results were judged according to economic theory, statistical characteristics, and out of sample predictive accuracy.

THEORETICAL MODELS

The static form of the derived demand for imports in country $i$ at time $t$, was:

$$
\ln(Q_{it}) = a_0 + a_1 \ln \left( \frac{P_{itq}}{P_{ito}} \right) + a_2 \ln(Y_{it}) + \epsilon_{it}
$$

(1)

where $Y_{it}$ is the gross national output, $P_{itq}$ and $P_{ito}$ are, respectively, indices of the import price of the particular forest product and the price of all other goods and services in the economy, and the residual term is $\epsilon_{it} = \alpha_i + u_{it}$, where $\alpha_i$ is an unobserved effect that varies between countries, and $u_{it}$ is a time-varying effect within the country.

The dynamic model, based on adaptive expectations/partial adjustment theory (Johnston 1984) was:

$$
\ln(Q_{it}) = a_0' + \gamma a_1 \ln \left( \frac{P_{itq}}{P_{ito}} \right) + \gamma a_2 \ln(Y_{it}) + (1-\gamma)(Q_{i,t-1}) + \epsilon_{it}'
$$

(2)

where $0 \leq \gamma \leq 1$ is the adjustment speed.
DATA

Data on annual imports and unit value of imports were obtained for 1970 to 1997 for 64 countries and the main forest products, from the Food and Agriculture Organization FAOStat database.

Import prices were computed from the import values reported by the FAO in nominal U.S. dollars. This price was converted to local currency using the local exchange rate from the World Development Indicators database (WDI), deflated with the GDP deflator (from the WDI), and then converted to international dollars to reflect purchasing power parity (World Bank 2003). The import demand shifter, \( Y \) was real GDP, expressed in international dollars.

MODEL ESTIMATION AND TESTING

Models (1) and (2) were estimated with data from 1970 to 1987. The following methods were used:

- **Pooled OLS**—Minimized the sum of squares of the residuals. The significance of the unobserved country effect indicated omitted-variable bias.
- **First Differencing**—Eliminated the unobserved country effect by differencing the variables over adjacent years (Wooldridge 2000).
- **Fixed Effects**—Removed the unobserved country effect by replacing variables by their distance to their mean within each country and over time.
- **Random Effects**—Assumed that the unobserved country effect varies randomly across countries. The serial correlation of the residuals was handled by calculating the feasible generalized least squares estimator (Wooldridge 2000) by maximum likelihood. We checked the consistency of the estimates with Hausman’s (1978) test.
- **Arellano and Bond’s (1991) Method**—Is intended to avoid inconsistent estimates due to the correlation of the lagged dependent variable with the residuals in model (2).

First order serial correlation was tested as in Wooldridge (2000), and robust standard errors were calculated with White’s (1980) method.

For each model, imports were forecast for each country and product from 1988 to 1997, conditional on imports in 1988, the elasticity estimates, and the observed prices and

<table>
<thead>
<tr>
<th>Product</th>
<th>Pooled OLS</th>
<th>Pooled First Differencing</th>
<th>Random Effects</th>
<th>Fixed Effects</th>
<th>Arellano-Bond</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roundwood</td>
<td>0.93</td>
<td>0.98</td>
<td>0.97</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td>Sawnwood</td>
<td>0.80</td>
<td>0.97</td>
<td>0.94</td>
<td>0.91</td>
<td>0.97</td>
</tr>
<tr>
<td>Plywood &amp; veneer</td>
<td>0.83</td>
<td>0.92</td>
<td>0.91</td>
<td>0.89</td>
<td>0.92</td>
</tr>
<tr>
<td>Particleboard</td>
<td>0.56</td>
<td>0.72</td>
<td>0.68</td>
<td>0.69</td>
<td>0.72</td>
</tr>
<tr>
<td>Fiberboard</td>
<td>0.74</td>
<td>0.72</td>
<td>0.75</td>
<td>0.75</td>
<td>0.72</td>
</tr>
<tr>
<td>Chemical pulp</td>
<td>0.74</td>
<td>0.72</td>
<td>0.75</td>
<td>0.75</td>
<td>0.72</td>
</tr>
<tr>
<td>Recovered paper</td>
<td>0.90</td>
<td>0.98</td>
<td>0.86</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td>Newsprint</td>
<td>0.73</td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>Printing and Writing paper</td>
<td>0.75</td>
<td>0.83</td>
<td>0.83</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>Other Paper &amp; Paperboard</td>
<td>0.87</td>
<td>0.92</td>
<td>0.92</td>
<td>0.91</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Coefficient of determination \( R^2 \) in regression of predicted imports on observed imports, across all countries and years. Bold characters indicate the best forecasts.
Table 2—Multi criteria comparison of static and dynamic models estimated with different methods.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Static Model</th>
<th>Dynamic Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pooled OLS</td>
<td>First Differencing</td>
</tr>
<tr>
<td>Correct sign</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Price</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Income</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Significance</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Goodness of fit</td>
<td>-10</td>
<td>-9</td>
</tr>
<tr>
<td>Autocorrelation &gt;0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Endogeneity</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Forecasting accuracy</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>33</td>
</tr>
</tbody>
</table>

*a* Number of products with the signs of price and income elasticities consistent with economic theory.

*b* Number of products with elasticities significantly different from zero at the 5% level.

*c* Number of products with the lowest within-sample RMSE.

*d* Number of products for which there was significant autocorrelation, and the autocorrelation was greater than 0.5.

*e* Number of products with explanatory variables correlated with the error term.

*f* Number of products for which the $R^2$ of the observed against the predicted imports was highest (table 1).
GDP from 1988 to 1997. The forecasting equation was based on the differential form of model (2):

\[ \ln(\hat{Q}_{i,t+1}) = \ln(\hat{Q}_{i,t}) + \hat{a}_1(\ln(P_{i,t+1}) - \ln(P_{i,t})) + \hat{a}_2(\ln(Y_{i,t+1}) - \ln(Y_{i,t})) + \hat{a}_3(\ln(\hat{Q}_t) - \ln(\hat{Q}_{t-1})) \]

where \( \hat{a}_3 = 0 \) for the static model.

The forecasting accuracy was measured by the coefficient of determination in a regression of observed on predicted imports, across all countries and years from 1988 to 1997.

**SUMMARY OF RESULTS**

**Long-Term Post-Sample Forecasting Errors**

The \( R^2 \)'s of actual on observed imports from 1988 to 1997 are in table 1. The dynamic model estimated by first differencing gave the best predictions for eight of the ten products. However, the static model estimated by first differencing gave predictions that were nearly as accurate for all products.

The worst predictions were obtained, for all products, with the dynamic model estimated by pooled OLS. Very poor predictions were also obtained for some products with the dynamic model estimated either with fixed or random effects.

**Multi-criteria evaluation**

To compare the different models and the estimation methods, we scored each model-method combination according to the criteria in table 2. Estimation of the static and dynamic models, by all methods gave the theoretically expected sign for the elasticities. In most cases the elasticities were significantly different from zero. The best fit was obtained with the dynamic model estimated by fixed effects. There was less serial correlation in the dynamic than in the static formulation. Endogeneity was a problem mostly for the dynamic model estimated by pooled OLS, first differencing, and fixed and random effects. The dynamic model estimated by first differencing was most accurate in post-sample forecasting.

The totals in the last row of table 2 give equal weight to each of the six criteria. However, if the particular interest in using these models was for predicting future import demand then forecasting accuracy would be weighted relatively more. Giving equal weight to each of the six criteria suggests that the dynamic formulation estimated with the Arellano-Bond method gave the best empirical models of the demand for forest product imports. The long-term elasticities implied by the Arellano-Bond method are in table 3. Most elasticities had the expected theoretical sign and were statistically significant. Imports were inelastic with respect to price for most products, except fiberboard, and printing and writing paper. For all products, imports were elastic with respect to income.

**ACKNOWLEDGMENTS**

The research was supported in parts by USDA_CSREES NRI grants 98-35400-6110 and 2003-3540-13816, by McIntire-Stennis grant 4456, and by the School of Natural Resources, University of Wisconsin-Madison. We thank two anonymous reviewers for their very useful comments on a previous draft of this paper.

---

Table 3—Long-term price and income elasticity of import demand from the dynamic model estimated with the Arellano-Bond method, 1970 to 1987.

<table>
<thead>
<tr>
<th>Commodity</th>
<th>Price</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roundwood</td>
<td>-0.74 (0.42)</td>
<td>2.21 (0.86)*</td>
</tr>
<tr>
<td>Sawnwood</td>
<td>-0.49 (0.18)**</td>
<td>2.71 (0.68)**</td>
</tr>
<tr>
<td>Plywood/ veneer</td>
<td>-0.81 (0.21)**</td>
<td>2.74 (0.58)**</td>
</tr>
<tr>
<td>Particleboard</td>
<td>-0.70 (0.29)*</td>
<td>5.70 (1.41)**</td>
</tr>
<tr>
<td>Fiberboard</td>
<td>-1.53 (0.31)**</td>
<td>1.76 (1.23)</td>
</tr>
<tr>
<td>Chemical Pulp</td>
<td>-0.48 (0.18)**</td>
<td>2.72 (0.62)**</td>
</tr>
<tr>
<td>Recovered Paper</td>
<td>0.01 (0.15)</td>
<td>2.50 (1.30)</td>
</tr>
<tr>
<td>Newsprint</td>
<td>-0.50 (0.22)*</td>
<td>1.13 (0.41)**</td>
</tr>
<tr>
<td>Printing and Writing Paper</td>
<td>-1.20 (0.32)**</td>
<td>1.47 (0.59)*</td>
</tr>
<tr>
<td>Other Paper and Paperboard</td>
<td>-0.74 (0.27)**</td>
<td>1.14 (0.34)**</td>
</tr>
</tbody>
</table>

*, ** Significant at 5% or 1% level, respectively.
LITERATURE CITED


This page is intentionally left blank.
A DEMAND-BASED SCENARIO OPTIMIZATION MODEL FOR SUPPLY OF RAW MATERIAL TO THE FOREST INDUSTRY

Daniel Hultqvist\(^1\), Leif Olsson\(^2\)

ABSTRACT

There are usually many sources for the supply of raw material to a pulp or paper mill in Sweden. The decision-maker can procure raw material from the company’s own forest(s), buy in from the owners of small forests, swap with other forest operators or purchase on the international market. Wood chips for these mills are brought from sawmills and have their own transportation system, but are included in the model presented here. Optimizing the supply of raw material at a tactical planning level is a challenging task, and can only be managed properly if elements of uncertainty are considered. The solution would otherwise give few options when disturbances occur, for example when weather conditions change. The deterministic equivalent of the scenario model is formulated and solved as a convex mixed integer quadratic model. In this paper, we have focused on model building since this is a challenging task in itself. The model is tested on one small sample problem and one full-scale problem. We show how to simulate and evaluate different strategies in the flow of raw materials, from harvesting operations to delivery at the mills. In our study, the full-scale problem has been solved within reasonable time using commercially available software, and the solution indicates that the hedging effect in the stochastic solution adds flexibility. Future work will involve the development of stochastic programming algorithms to solve large-scale instances of the problem using parallel computing.

KEYWORDS: Stochastic programming, forest logistics, flexibility, load-bearing capacity, mixed-integer quadratic programming.

INTRODUCTION

In recent years, the forest companies in Sweden have moved their focus from harvest decisions, which have been relatively well investigated, to forest logistics (Karlsson and others 2002, and Karlsson 2002) and especially road investments (Olsson 2005; Olsson 2003 and Arvidsson and others 2000), and storage of roundwood (Persson and others 2002; Persson & Elowsson. 2002). Case studies have indicated that optimal solutions for the road investment problem can be calculated in an efficient manner, as shown by Olsson (2003 and 2005) and by Arvidsson and others (2000).

As opposed to road investments, there are no definitive figures for the cost of roundwood storage. Although studies exist, such as those by Persson and others (2002) and Persson & Elowsson (2002), it is in general very hard to estimate the cost of loss in value. It is also difficult to estimate how roundwood of low quality will affect the quality of the end product at different mills. The problem will vary depending on the assortment, weather conditions and the
industrial process, as well as storage location etc. Hence, only “rule of thumb” estimations of the costs are used here. Furthermore, it cannot be assumed that decisions involving harvesting, transport and the forest industry are independent from storage decisions, as first put forward by Barros & Weintrub (1982) and more recently by, for instance, Karlsson (2002). There are also studies that indicate substantial gains if the process at the mills is integrated with the planning of the raw material procurement (Arlinger and others 2000).

Here we describe and test an optimization model that integrates different aspects of the raw material supply chain from harvesting to industrial processing. The model focuses on procurement of raw material from forest owned by the mill. We have, however, included the purchase of roundwood on the domestic and international market in the model.

For sawmills, the raw material supply is a pure roundwood supply chain problem. This is not the case for pulp or paper mills, however, where raw material procurement also includes the supply of wood chips. This is transported from the sawmills to the pulp and paper mills on special trucks. This approach makes integrated planning an important task for the staff at forest companies, and is also the main thrust of our integrated optimization model presented here. Nonetheless, the supply of wood chips is rather easily managed, since it only considers transportation from sawmills to pulp or paper mills by main roads. Hence, this paper will concentrate on activities involving the roundwood supply chain.

In contrast to other attempts to optimize the raw material supply with purely deterministic models as in, for example Karlsson (2002), our stochastic model focuses on procurement with respect to uncertain load-bearing capacity of roads and harvesting grounds.

We have developed a scenario optimization model (Birge & Louveaux 1997; Kall & Wallace 1994) with dynamic information constraints (Lohmander & Olsson 2003). The deterministic equivalent is directly implemented as a convex mixed integer quadratic model (Wolsey 1997).

We describe the features of the model using a sample problem and show that it is possible to solve a full-scale stochastic model with commercial software (Anon. 2002) and a standard PC to near optimality within a reasonable length of time, with a limited set of scenarios and periods. Nevertheless, later our intention is to use decomposition techniques to solve the problem in parallel. Useful decomposition algorithms exist for this type of problem as described by, for instance, Linderoth & Wright (2003), and affords the opportunity of solving more complex models.

Cost estimations, detailed mathematical formulation and detailed problem definition have not been included here, since space is limited. However, a much more detailed presentation of this work can be found in Hultqvist & Olsson (2004).

An outline of the remainder of this paper follows.

We start with an overall description of the raw material supply problem being considered, together with our assumptions. In section three, we present a verbal description of the mixed integer quadratic model. In section four, we present a sample problem and a full-scaled problem. In section five, we present results from the optimization of a sample problem, and also some results from the modelling of a full-scale problem. The last section contains the discussion and concluding remarks.

OVERVIEW OF THE PROBLEM

We only present an overview of the raw material supply chain problem here, and solve the scenario model for a period of six months. A detailed description of all component parts can be found in Hultqvist & Olsson (2004). The activities in the roundwood supply chain are described below.

- The first step is to harvest the trees with the harvester and put the trees in log piles.
- Then the log piles are picked up with a forwarder and transported to a roadside storage area. We assume that there are no costs related to roundwood storage in the actual forest, since good planning will keep forest storage at a low level. If using a harwarder, a combined machine which is both a harvester and a forwarder, this problem does not even exist.
- We also have the opportunity to buy roundwood from small forest-owners and pick it up at the roadside.
- Stumpage is included as additional harvest sites among the company’s own forest. In Sweden, these sites must be harvested within three years.
- It is possible to harvest sites that are not accessible during the spring thaw in advance. The harvested roundwood is then hauled to a road that will be accessible during the thawing period, and stored at the roadside until required. This form of pre-hauling is common in Sweden.
- We then haul the roundwood on the roads to a storage terminal, a better road or directly to the mills.
If it is not possible, or too expensive, to procure a sufficient volume of roundwood for the mills from the company's own forest, we must purchase wood on the roundwood market or, to solve short-term problems, use the security stock.

Infrastructure investments are not considered in the model since these are long-term decisions, as described in Olsson (2003) and Olsson (2005).

The model can handle railway transportation.

The other part of the raw material procurement for pulp and paper mills involves wood chips.

Wood chips for pulp and paper mills are procured from sawmills, and have their own transportation system. Haulage of wood chips is done on roads that are always accessible.

The description above has of course been simplified to give an overall impression, and more details are given in Hultqvist & Olsson (2004). However, one essential part of the model is that road transportation of the roundwood is affected by the load-bearing capacity of the roads, as described by Hossein (1999). The variation in load-bearing capacity of a road depends mostly on weather phenomena, and must be modeled as uncertain. In the model presented here we use scenarios, also called events, as depicted in Fig. 1 to include uncertainties in load-bearing capacity of the roads and the ground. This approach is well described in Lohmander & Olsson (2003), and is the common way of modeling uncertainty in the stochastic programming community (Birge & Louveaux 1997; Kall & Wallace 1994).

The event tree (Fig. 1) considers uncertainty in load-bearing capacity of the ground and roads. These stochastic events have the following definitions:

- L = Low water content in the ground, keeping more roads and harvest areas open.
- M = Normal water content in the ground, keeping a normal number of roads and harvest areas open.
- H = High water content in the ground, causing more roads to be closed and making more harvest areas than usual unavailable.

We assume that it is equally likely that there will be dry, wet or normal weather conditions in our general model. However, in reality, the scenario distribution can be estimated from historical weather data, for any geographical area, and generated with moment-matching scenario generation, for example (Höyland and others 2003). Our main point here is, however, that even a uniform distribution includes more flexibility in the solutions than a pure deterministic model. This flexibility gives solutions that hedge against future disturbances (Wallace 2000 and Wallace 2003). This general hedging effect is the main reason why we use a stochastic model here.

**MATHEMATICAL MODEL**

The mathematical model is only described in words here, since a full description would take up too much space. A detailed mathematical description of the convex mixed-integer quadratic model can be found in Hultqvist & Olsson (2004).

**The objective function**

The purpose of the objective function is to minimize the cost of supplying the industries with the volumes demanded. The costs included are described below.

Harvest:

- I) The time-dependent cost of harvesting.
- II) The fixed cost for harvesting machines.
- III) The cost of transporting the crew to and from the harvest site.

Transportation and roads:

- IV) The initial cost of harvesting at a site. This includes transportation of equipment and some maintenance of roads.
VI) The volume-dependent haulage cost.
VII) The cost of pre-hauling to accessible road(s) during the thawing period.

Storage:
VIII) The volume-dependent storage cost.
IX) The cost of unloading and measuring incoming volumes at storage terminals and mills.
X) The cost of loading outgoing volumes at the terminals.
XI) The fixed cost of having and using storage terminals.

Markets:
XII) The cost of buying roundwood on the open domestic market.
XIII) The collection cost for roundwood purchased on the domestic market.
XIV) The cost of buying roundwood on the international market.

Sawmills:
XV) The extra cost of having more than one assortment at a sawmill.

Railroads:
XVI) The fixed cost for operating a railroad system.
XVII) The volume-dependent cost of railroad transportation, including loading and unloading from truck to train.

Wood chips:
XVIII) The transportation and purchase cost of wood chips from saw mills to pulp or paper mills.

Constraints
There are many model constraints, described below, which must be considered for the raw material supply problem.

Harvest:
1) The relative part harvested in each time period and in each scenario must be less than one. This constraint sets a binary variable that indicates whether anything is harvested during a specific time period and scenario.

2) The sum of the parts harvested in all time periods, for each harvest site and for each path through the scenario tree, cannot be greater than the total volume available.

3) The number of time periods within which a specific area can be harvested is restricted to two. This is done to simplify the model and to get better behaviour from the solution. All harvest sites can be harvested in one time period or less. If harvesting of a site starts at the end of one time period, it will be allowed to continue into the next period, but there is no reason for it to continue into a third period. This would only lead to a waste of capacity, to have machines standing still at that site. It will also set the binary variable that indicates whether some volume is harvested in the area.

4) If an area is harvested during two consecutive time periods, a binary variable is set to one to ensure that no start-up cost is charged for the second time period.

5) The number of areas that are harvested in two consecutive time periods is restricted to the actual number of harvesting teams available.

6) This constraint is constructed to indicate whether an area is being harvested during two non-consecutive time periods. If this is the case, a binary variable is set to give the extra cost this implies.

7) All the activities that a harvesting team is involved in are converted into hours so that it will be possible to include time restrictions in the model. This covers both harvest and transfer times. If the next time period is the thawing period, harvest hours for this period are added to the previous period. The reason for this is that areas with harvest volumes needed during thawing are harvested during the winter and pre-hauled to a location that is accessible under the thawing period.

8) The number of harvest hours can be restricted, both with respect to the minimum and the maximum number of available hours. A binary variable for each time period is set if any harvesting take place. This gives the fixed harvesting costs for machines, set at II in the objective function.

9) For each assortment, the difference between outgoing and ingoing flow of roundwood at a harvest node must equal the volume harvested at the node, since the whole of the harvested volume should be transported from the harvest site.

Transportation and roads:
10) At a crossroads, the ingoing and outgoing volume of roundwood must be the same for each assortment, since this represents transhipment nodes in the network.

11) The total volume transported on a road can be restricted.
This is useful on small roads that cannot handle much traffic in an appropriate way.

12) The number of hours that can be used for transport of roundwood can be restricted. This, then, would include time for loading, unloading and driving.
Storage:
13) At the first time period, initial storage volume and volume transported into the storage must equal what is transported out and stored for later use.
14) Stored volumes from earlier time periods and the volume transported for storage must equal storage for later use or the volume transported out from the storage.
15) In the final time period, stored volume and the volume transported for storage minus the volume transported out of storage must be greater than or equal to a fixed volume. This volume is set in advance, and represents the final minimum storage level.
16) The total volume stored must be less than or equal to the maximum storage capacity at each storage node.
17) For each assortment in storage, a minimum volume (safety level) can be set. This forces the model to always keep a given amount of an assortment in storage.

Mills:
18) The volume delivered to mills from other owners can have a maximum and a minimum restriction. The volume transported to a mill, the volume of roundwood bought on the open market and the volume delivered from other mill owners due to swapping is summed into a continuous variable. The model can thereby decide the cheapest way of supplying the mill with roundwood of different assortments from different sources.
19) The volume transported to a mill, bought on the open market and delivered from other mill owners due to swapping, is also summed into a continuous variable as described above.
20) The sum of all assortments must equal the total demand at the mill for each time period. This condition ensures that the total demand volume is correct, even if the sum of demand for all the different assortments is less.
21) For each assortment except one used by a saw mill, an extra location cost is added.
22) For each time period and assortment, the volume delivered at a mill is not allowed to be less than the minimum volume demand at that mill.

Wood chips:
23) At sawmills, some portion of the demand will be returned into the system as wood chips to be delivered to pulp or paper mills. This portion can be specified here.

24) The volume of wood chips delivered at a pulp or paper mill must meet the demand.

Markets:
25) There is a limitation as to how much roundwood can be bought from the domestic market. This is due to the fact that the domestic market is usually limited, or that it might be impractical to buy roundwood from too far away.
26) Even when buying on the open international market, there might be limitations on how much can be bought. These restrictions are related to transport capacity, or exist for political reasons.

Railroads:
27) The volume transported into a railroad terminal by trucks is transported out of the same terminal by train.
28) There is a volume limitation as to how much roundwood can be transported by a railroad during each time period. When a railroad system is used, a binary variable is set to give the fixed cost of using a railroad system.

Variables:
29) All continuous variables are equal to or greater than zero in the model.

Problems
We have tested the model on the sample problem depicted in Fig.2, as well as on a full-scale problem (Fig. 4). The sample problem is as small as it can be and still consist of all the interesting parts in the supply chain, in other words harvest areas, roads of different classes, a storage terminal, a saw mill and a pulp or paper mill.

These roads have different classes, depending on the construction of the road (Fig. 3). The paved roads are all A-class roads. The classes B through D are all gravel roads. Their accessibility is described in Table 1 and has been taken from Löfroth and others (2000).

The real case problem depicted below involves 450 roads and 320 harvest areas. The road network covers an area of about 11,500 km² in the central part of Sweden (Fig 4).

RESULTS
All the results given below have been computed on a 2.0 GHz Pentium 4 desktop computer with 512 Mb internal memory. The program used for the calculations was Lingo
Figure 2—A sample problem depicted as a network, with 18 possible harvest areas and 20 roads. The network has one storage location, one sawmill and one pulp/paper mill.

Figure 3—The classifications for each of the 20 roads in the sample problem.
8.0 from Lindo Systems (Anon 2002). The difference in the solution time for the stochastic and deterministic versions of the model applied to the sample problem was small (Table 2).

As seen in Table 2, the differences in solution time were substantial for the larger real case problem.

For the large real case, an optimal solution for the stochastic model could not be found within reasonable time. A solution less than 0.1 per cent from global optimum was found within a minute of calculation. For the deterministic problem, the same level of near-optimum solution was also found within a minute of calculation. Even if the parameters of the problems were changed to give a first solution almost 20 per cent from global optimum, a solution less then 0.5 per cent from global optimum was found within a minute or two of calculation. Before the actual calculations start, the program generates and scales the matrices. This takes just a few seconds for a small problem. However, for a large stochastic problem this takes over one hour. (Table 3)

For the sample problem, the same eight harvest sites were harvested using both the stochastic model and all scenarios with the deterministic model. These sites can be seen in Figure 5.

Even though the same eight harvest areas are chosen to be harvested during the planning horizon, there are some differences between the stochastic and the deterministic approaches. These differences are related to when any particular site is picked to be harvested in time. For the large problem it is mainly the same areas that are harvested, but there are exceptions. Some harvest areas found in the deterministic solutions are not in the stochastic solution, and the other way around. There are also the same differences between the two types of solution as found in the sample problem, in other words when an area is to be harvested in time.

### Table 1—Times when the roads of different classes are open or closed. The meaning of the notation for different scenarios is described above.

<table>
<thead>
<tr>
<th>Road class</th>
<th>Winter</th>
<th>Spring Low</th>
<th>Spring Medium</th>
<th>Spring High</th>
<th>Summer Low</th>
<th>Summer Medium</th>
<th>Summer High</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Open</td>
<td>Open</td>
<td>Open</td>
<td>Open</td>
<td>Open</td>
<td>Open</td>
<td>Open</td>
</tr>
<tr>
<td>B</td>
<td>Open</td>
<td>Closed</td>
<td>Closed</td>
<td>Closed</td>
<td>Open</td>
<td>Open</td>
<td>Open</td>
</tr>
<tr>
<td>C</td>
<td>Open</td>
<td>Closed</td>
<td>Closed</td>
<td>Closed</td>
<td>Open</td>
<td>Open</td>
<td>Closed</td>
</tr>
<tr>
<td>D</td>
<td>Open</td>
<td>Closed</td>
<td>Closed</td>
<td>Closed</td>
<td>Open</td>
<td>Closed</td>
<td>Closed</td>
</tr>
</tbody>
</table>

### Table 2—Size and solution time for the stochastic and deterministic model formulations of the sample problem.

<table>
<thead>
<tr>
<th></th>
<th>Stochastic Programming</th>
<th>Deterministic Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
<td>2 311</td>
<td>626</td>
</tr>
<tr>
<td>Integer variables</td>
<td>344</td>
<td>153</td>
</tr>
<tr>
<td>Quadratic variables</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Constraints</td>
<td>2 927</td>
<td>661</td>
</tr>
<tr>
<td>Iterations</td>
<td>12 486</td>
<td>4 000 - 40 000</td>
</tr>
<tr>
<td>Solution time</td>
<td>54 seconds</td>
<td>5-40 seconds</td>
</tr>
</tbody>
</table>
DISCUSSION

The Small Sample Problem

The difference between the stochastic and deterministic solutions for the sample problem is mainly a question of when to harvest. The areas to harvest in both types of problem are the same, but they are not harvested in the same order. The reason that the same areas are harvested is probably due to the very small problem size and to the fact that a large proportion of the available volume of roundwood is harvested. This limits the number of solutions available. The sites chosen are closest to the mills and usually have good ground and bearing capacity of roads, in other words accessibility. Not all of the selected harvest sites are available during all time periods. Hence, for some areas the harvested volume has to be pre-hauled. This comes with an extra cost of transportation, making it a more expensive way of supplying the mills. Pre-hauling occurred in the stochastic solution and in five of the nine deterministic solutions.

Table 3—Size and solution time for the large real case problem with stochastic and deterministic model formulation. The global optimal solution for the stochastic programming problem could not be found within a reasonable time frame. The solution obtained after 1 hour of calculation is not improved, more then very marginally, even when the computer was left running for closer to a week.

<table>
<thead>
<tr>
<th></th>
<th>Stochastic Programming</th>
<th>Deterministic Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
<td>39 363</td>
<td>10 353</td>
</tr>
<tr>
<td>Integer variables</td>
<td>5 779</td>
<td>2 569</td>
</tr>
<tr>
<td>Quadratic variables</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Constraints</td>
<td>48 883</td>
<td>10 417</td>
</tr>
<tr>
<td>Iterations</td>
<td>—</td>
<td>50 000 - 850 000</td>
</tr>
<tr>
<td>Generation and scaling time</td>
<td>1 hour 6 minutes</td>
<td>3 minutes 42 seconds</td>
</tr>
<tr>
<td>Calculation time to near-optimum solution</td>
<td>1 minute</td>
<td>0.5 – 2 minutes</td>
</tr>
<tr>
<td>Calculation time to global optimum solution</td>
<td>—</td>
<td>0.25 – 5 hours</td>
</tr>
</tbody>
</table>

Figure 5—The sites harvested in the sample problem.
The Large Real Case Problem

For the large problem, the stochastic and deterministic models do not always choose the same harvest areas. The solutions are generally the same, but there are areas in the stochastic solution that are not in the deterministic solution. There are also areas in the deterministic solution that are not present in the stochastic solution. For those that are the same, there are (just as for the sample problem) differences in when they are harvested. The fact that there are harvest sites chosen in the stochastic solution that are not even present in the solutions for the deterministic problem indicates that uncertainty in load-bearing capacity of roads and harvest sites does matter. One must remember that the stochastic programming solution hedges against uncertainty, making the model generate more stable solutions.

One important problem with the stochastic formulation is that it is not possible to solve it to a global optimum within a reasonable time. What has been called the real case above is only part of the real-world supply chain. It represents one out of five districts supplying a pulp mill in the center of Sweden. Hence, if we were to use this model to solve the whole supply chain, we would need an alternative approach to making use of a desktop computer. One approach would be to use a cluster of computers to solve smaller parts of the problem. Algorithms are available for this, which solves the stochastic problem by solving many deterministic equivalences in each iteration. This would make it possible to solve larger problems much more quickly and with a higher degree of detail in the model.

ACKNOWLEDGEMENTS

We wish to thank the following for financial support, without which this research would not have been possible: FSCN at Mid-Sweden University, the Computer Science Department at Umeå University, the Kempe Foundation, NorFa and the Gunnar Sundblad Research Fund. We also wish to thank and express our gratitude to numerous people at Holmen Skog for answering forestry-related questions, and for providing the dataset. Without the latter, we would not have been able to test our model on a real-world problem. Lastly, we thank Prof. Peter Lohmander at the Swedish University of Agricultural Sciences for many useful discussions, and also advice.

LITERATURE CITED


Löfroth, C; Adolfsson, J; Persson, B; Persson, T; Bergström, J. 2000. Handledning av NVDB-registering av vägdata, skogsbilvägar. The Forest Research Institute of Sweden.


Wallace, S.W. 2003. Decision making under uncertainty: the art of modelling. Molde University Collage 15.06.03.

ANALYSIS OF AN EXTENSIVE THINNING PROGRAM ON EASTERN OREGON NATIONAL FORESTS BY USING A DYNAMIC SPATIAL EQUILIBRIUM MARKET MODEL WITH ENDOGENOUS INDUSTRY CAPACITY

Greg Latta\(^1\) and Darius Adams\(^2\)

ABSTRACT

An intertemporal, spatial equilibrium model of the eastern Oregon forest products sector was employed to estimate the impacts of a thinning program to help restore ecosystem health on national forests in the region. Harvest and silvicultural decisions of private landowners and the output and capacity investment decisions of mills were endogenous. Thinning treatment areas were considered in the spatial context of the model, but only thinning areas with non-zero board foot volume were included. Simulations suggested that 63% of the treatable area and 91% of the board foot volume in thinning areas could be removed in a 20-year program with no stumpage charges or hazard removal costs. These volumes would replace declining private harvests in the first decade but would not reduce the projected long-term drop in regional cut. Projected reductions in mill numbers would be postponed until after 2030.

INTRODUCTION

The eastern Oregon forest products sector is an industry in decline. Precipitated in part by the near elimination of national forest harvest, the sector has lost more than half of its processing capacity since the early 1990’s. Based on private timber resources alone, further reductions in harvest seem likely (Adams and Latta, 2003). At the same time, wide-scale thinning programs have been proposed to reduce volumes in overstocked stands on national forests in the face of growing concerns about forest health, fire hazards and extreme fire behavior\(^3\). This paper employs a dynamic, spatial equilibrium model of the eastern Oregon log market to examine the potential harvest, price and milling capacity impacts of such programs.

A MODEL OF THE EASTERN OREGON FOREST SECTOR

The model developed here builds on the work by Krumland and McKillop (1990) and Bare and others (1995) in identifying timber supply/harvest activities on private lands at the finest possible geographic scale for a range of silvicultural options. This supply module is coupled with a representation of demand in an intertemporal optimization model of the eastern Oregon softwood log market following work by Adams and others (1996), Schillinger and others (2003) and Adams and Latta (2003). Total harvest, market prices, flows from forests to mills, and volumes processed by mills are endogenous, as are silvicultural investment decisions on private lands. Transportation costs between timber harvest sources and log processing centers are explicit. Decisions to maintain, expand, contract or close processing capacity are also endogenous based on costs and profit potential. National forest thinning options are introduced as additional harvest sources. The analysis provides estimates of the effects of a thinning program on mill locations and processing capacity and of the proportion of national forest thinning opportunities that could be economically undertaken with existing mill technology.

\(^1\) Greg Latta is a faculty research assistant and \(^2\) Darius Adams is a professor in the Department of Forest Resources, Oregon State University, Corvallis, OR 97331 USA.
\(^3\) The Healthy Forests Restoration Act of 2003 authorizes such programs on public lands.
Because data limitations prevent estimation of demand directly at the mill level, we employ a simple proportioning procedure to disaggregate regional demand estimates to individual mills. This demand includes a processing capacity variable (representing quasi-fixed capital stock). Output is bounded above by capacity and below by a minimum economic level of capacity utilization. Capacity can be maintained or expanded in a three-level cost structure (maintain, replace depreciation, expand) or left idle to deteriorate. Firms decide on their optimal intertemporal capacity strategy based on the present value of future costs and returns.4

Log Supply

Inventory data for this study derive from the Forest Service’s 1999 remeasurement of permanent plots on private forestland in eastern Oregon. Our work employed a preliminary version of this inventory, which differs somewhat from the draft final release [see Azuma and others (2002) for a description of inventory methods and broad results]. The primary differences lie in recomputation of some site index values, assignment of vegetation type for some plots and expansion factors. In this inventory, plots are divided into “condition classes”, which are portions of the plot comprising uniform vegetation type, stand size, stocking and previous harvest treatment. Data cover 529 condition classes on 492 plots on timberland and other forestland. We treat the condition class as the basic resource unit with separate yield projections for each class.

Management practices were divided into reserve (no harvest), even and uneven-aged classes with three increasing levels of intensity in non-reserve groups, termed management intensity classes (MIC), as shown in Table 1. These regimes were adapted from work by Bare and others (1995) in an analysis of harvest potential in roughly comparable forest types in eastern Washington. They are expressed in threshold form. That is, stands must meet minimum stocking conditions before harvest can occur and harvests must exceed minimum removal volumes. Allocation of land to the uneven-aged option can occur only at the start of the projection and is irrevocable (there is no switching between uneven and even-aged MICs). As a consequence, several timing options for taking the first partial cut in uneven-aged regimes were introduced in the activities employed in the optimization (stand volume could rise above minimum thresholds to varying degrees before the first cut in the projection was made).

Log Demand

Demand Equations. Mill locations, approximate output, and capacity were derived from industry directories [see, for example, Random Lengths, 2003]. During the first projection period (1996-2000) the sector comprised 15 lumber mills at 13 locations or processing centers in the

4 This view of the capacity decision is similar to that employed in so-called “third generation” econometric profit function models with endogenous capital stock. Firms are seen as choosing their capital stock trajectories (simultaneously with output) so as to maximize the firm’s present net worth. See, for example, Stevens (1995).
Data are not available to estimate log demand and other production characteristics at the mill or processing center level. As a consequence we estimate log demand for the regional industry then disaggregate these relations to the mills using approximate methods.

Regional log demand was estimated using a normalized, restricted quadratic profit function. The industry was assumed to have one output (lumber), with residues treated as by-products. Inputs include three categories: logs, labor and other variable inputs. Capital stock (measured here as physical processing capacity) is treated as quasi-fixed and technology is represented by a time trend. The industries are assumed to be competitive, attempting to maximize profits subject to endogenous prices of logs and exogenous prices of output, labor and other variable inputs. Since products from the region compete in international markets and represent a small share of overall market volume, treatment of output price as exogenous seems justified.

The empirical model consisted of the log demand equation together with the output supply, labor demand and profit function equations. Normally distributed stochastic disturbances with zero mean and constant variance were appended to each equation. Dummy variables were included to represent the dramatic effects of recession in the years 1980-1982. Time series data with annual observations from 1970 to 1999 were used in the estimation. Descriptions of data development and sources are given in Adams and Latta (2003).

Since virtually all of the data series exhibited some evidence of nonstationarity, coefficients were estimated with data in first difference form using iterative nonlinear three-stage least squares. The instrument set included current and lagged values of exogenous variables and lagged values of the endogenous variables. Symmetry was assumed and global convexity was imposed as described by Wiley and others (1973). Parameter estimates, asymptotic t-ratios and goodness-of-fit statistics for the log demand equation are given in table 2. These parameter estimates yield unconditional (Marshallian) own-price elasticities of wood demand for lumber production of -.29 (at sample means). Breusch-Godfrey tests on residuals of the log demand equation lead to rejection of the hypothesis of 1st and 2nd order autocorrelation at the .01 level.

Lacking time series data on mill inputs, outputs and costs, we estimated mill log demands by scaling the regional demand equation to the mill level using estimated mill outputs in 1998. This amounts to assuming that log demand at each mill has the same elasticities for log price and other demand variables as the regional equation. Similar assumptions have been used in other studies to derive demand elasticities at finer geographic scales [see, for example, Kutcher (1983) and Adams and Haynes (1996)], but this is a relatively restricted approach.

Capacity constraints. Additional relations were added to the model to allow mill capacity to change over time, shifting both the log demand equation and the capacity

---

Table 1—Silvicultural regimes (management intensity classes or MICs) for non-reserve groups used in the eastern Oregon analysis.

<table>
<thead>
<tr>
<th>EVEN-AGE</th>
<th>UNEVEN-AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clearcut if stand volume at least 12 MBF/acre residual. Natural regeneration.</td>
<td>Cut if stand volume at least 7 MBF/acre, leaving 4 MBF/acre in trees 7 inches and larger.</td>
</tr>
<tr>
<td>Clearcut if stand volume at least 13 MBF/acre residual. Plant to 250 trees/acre.</td>
<td>Cut if stand volume at least 9 MBF/acre, leaving 5 MBF/acre in trees 7 inches and larger.</td>
</tr>
<tr>
<td>Clearcut if stand volume at least 16 MBF/acre. Plant to 250 trees/acre, and thin to 175 trees/acre when stand height at least 15 feet.</td>
<td>Cut if stand volume at least 9 MBF/acre, leaving 4 MBF/acre residual in trees 7 inches and larger, underplant 100-150 trees per acre.</td>
</tr>
</tbody>
</table>

MBF = thousand board feet

---

5 The one remaining plywood mill in the region closes in the fourth period of our projection and does not reopen.

6 Eastern Oregon lumber output was 1.6% of total US consumption in 1999 and 2.0% in 2002.
bounds on log demand. Since we only consider log demand (and not product output) in the model, a mill’s physical production capacity is reflected in an upper bound on log demand (see figure 1). Over time, capacity follows the usual inventory identity: capacity this period is equal to capacity last period less depreciation plus investment. Additional capacity can be purchased in a three-tiered cost structure. Basic maintenance is required on all capacity and a low level of cost is charged for this maintenance so long as output is positive. This maintenance does not offset depreciation losses but is required to keep the mill running. A second tier involves repair and replacement of depreciated equipment. These costs do not expand capacity but simply maintain the current maximum output level. A third tier involves investment in new capacity to expand output beyond current levels. This is the most costly form of capacity and reflects new production lines or the costs of returning capacity to operational levels some time after a mill has closed.

Two additional constraints are placed on mills’ capacity changes and output to mimic observed behavior. First, a mill’s operating rate (output/capacity) can not fall below a preset minimum level unless the mill shuts down (output goes to zero). This is shown as the discontinuity in the log demand curve in Figure 1 at Q_M. This level of log use would correspond to the point of minimum average variable cost on the output marginal cost curve in traditional cost curve analysis. In the present case, as log price rises up the log demand curve the marginal cost of output would shift up. At a constant output price, minimum average variable cost would eventually meet, then exceed, price. The firm’s optimal strategy is to shut down, rather than lose more than its variable costs, and log demand drops to zero. We approximate this shut down-start up point with a minimum operating rate.

A second restriction sets a minimum permissible capacity size for each mill. This constraint posits that there is some minimum mill size below which it is not economic to operate. To be profitable and competitive, a mill must be at least some minimum size. In effect this threshold establishes a barrier to entry, or in the present case re-entry, to the regional industry. If a mill drops out of production and its capacity is allowed to depreciate, it must invest to bring capacity at least to the minimum threshold before restarting production.

Overall Model Structure
The model can be characterized as an optimizing, intertemporal, price endogenous, spatial equilibrium market model. Projections are made over some period (typically 100 years or more) using 5-year time intervals (years in figure are interval midpoints). Solutions are found by maximizing the present value of the willingness-to-pay integral under the demand curves less costs of timber growing, log

Table 2—Iterative 3SLS estimates of log demand equation coefficients for lumber mills with elasticities (at sample means) and associated asymptotic standard errors.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Elasticity</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wood Price</td>
<td>-.1266</td>
<td>.0420</td>
<td>-.287</td>
<td>.095</td>
</tr>
<tr>
<td>Output Price</td>
<td>.1401</td>
<td>.0494</td>
<td>.465</td>
<td>.164</td>
</tr>
<tr>
<td>Labor Price</td>
<td>-.5354</td>
<td>.3619</td>
<td>-.112</td>
<td>.076</td>
</tr>
<tr>
<td>Capacity_{t-1}</td>
<td>.6238</td>
<td>.1633</td>
<td>.879</td>
<td>.230</td>
</tr>
<tr>
<td>Trend</td>
<td>.3562x10^{-3}</td>
<td>.7830x10^{-3}</td>
<td>-.005</td>
<td>.011</td>
</tr>
<tr>
<td>Dummy 1980-82</td>
<td>.2537</td>
<td>.0491</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Equation estimated with variables as first differences. Squared simple correlation of actual and predicted differenced values 0.661.
processing and handling (including capacity investment) over the projection period. Results reported here use a 6% discount rate. Logs flow from condition classes (denoted by the subscript “CC” below) to processing facilities (the subscript “MILLS”). Capacity acts both to shift the log demand (willingness to pay) function and to bound the level of log demand (termed “receipts”) at each mill. The approach to timber inventory accounting might be viewed as a combination of Johnson and Scheurman’s (1976) model I and II forms. Even-aged stands are classified in part by dates of origin and next harvest, uneven-aged stands by the specific regime (MIC) in which they are enrolled.

The general structure of the model is outlined below (the time subscript is suppressed, except where essential, to simplify the notation):

1. \[ \text{MAX} \sum_{\text{TIME}} \left[ \sum_{\text{MILLS}} \text{Willingness to pay} \right. \]
   \[ = \left( \text{Receipts}_{\text{MILLS}}, \text{Capacity}_{\text{EOR}}^{(i-1)} \right) \]
   \[ - \text{Capacity costs (Maintenance, Depreciation, Expansion)} \]
   \[ - \text{Transport costs (CC \rightarrow MILLS)} \]
   \[ - \text{Harvest costs} \]
   \[ - \text{Planting and silvicultural costs} \right] (1 + r)^{\text{TIME}} \]
   \[ \text{+ Discounted Terminal Inventory Contribution} \]

Subject to:

2. All CCs must be allocated to some MIC in each period
3. Even-aged planting \( \leq \) area harvested in new and existing even-aged stands this period
4. \( \text{Harvest}_{\text{CC}} = (\text{Final Harvests + Thinnings})_{\text{EVEN-AGED AREAS}} + (\text{Partial Cuts})_{\text{UNEVEN-AGED AREAS}} \)
5. \( \text{Harvest}_{\text{CC}} \geq \sum_{\text{MILLS}} \text{Shipments}_{\text{CC, MILLS}} + \text{Exports}_{\text{CC}} \)
6. \( \sum_{\text{CC}} \text{Shipments}_{\text{CC, MILLS}} + \text{Imports} + \text{Public} \geq \text{Receipts}_{\text{MILLS}} \)
7. \( \text{Receipts}_{\text{MILLS}} \leq \text{Capacity}_{\text{MILLS}} \)
8. \( \text{Receipts}_{\text{MILLS}} \geq \mu \text{Capacity}_{\text{MILLS}} \text{ or zero.} \)
9. \( \text{Capacity}_{\text{MILLS}} = \text{Capacity}_{\text{MILLS}, \text{TIME}-1} (1-\sigma)^5 + \text{Maintenance}_{\text{MILLS}} + \text{Expansion}_{\text{MILLS}} \)
10. \( \text{Capacity}_{\text{MILLS}} \geq \xi \text{Capacity}_{\text{MILLS}, \text{INITIAL PERIOD}} \)
11. \( \text{Capacity}_{\text{EOR}}^{(i)} = \sum_{\text{MILLS}} \text{Capacity}_{\text{MILLS}}^{(i)} \)
12. Convexity Constraints (piece-wise linearization of area under demand functions)
13. Non-negativity

where

- \( \text{Capacity}_{\text{EOR}} \) is total eastern Oregon log processing capacity, as defined in (11),
- \( \text{CC} \) subscript refers to the basic condition class units that make up the inventory,
- \( \text{MILLS} \) subscript refers to the individual mills or log processing units,
- (i) and (i-1) superscripts refer to the solution iteration number (see discussion below),
- \( r \) is the discount rate,
- \( \sigma \) is the capacity depreciation rate,
- \( \mu \) is the minimum operating rate, and
- \( \xi \) is the minimum ratio of current to initial capacity (minimum plant size).

In the objective function (1), willingness-to-pay, the integral of the area under the demand curves, is a function of both log receipts and regional capacity (both endogenous). Constraints (2) and (3) are the standard accounting requirements for Johnson and Scheurman’s (1976) models I and II. Harvests (4) for each condition class in each period are the sum of final harvests and thinnings from even-aged areas plus partial cuts from uneven-aged areas. Each condition class in the sample represents a specific area in the inventory (determined by the class’s area expansion factor from the sample). In the solution, the area represented by each condition class can be broken down into a number of even and uneven-aged treatments that may vary over time. The logs harvested from each plot may be shipped to mills within the region or to destinations outside the region (5). Exports have accounted for about 25% of total eastern Oregon harvest and flow to an array of destinations in western Oregon. They are treated as exogenous in this analysis at a fixed 25% of total cut. Receipts at mills (6) comprise intra-regional log shipments from private lands, plus receipts from public lands, plus imports from outside the region. Imports are small and are treated as exogenous. Constraints (7)-(11) encompass the capacity model discussed above: receipts (mill output measured in units of log input) must be less than capacity (7) and greater than the minimum operating rate (8), capacity evolves over time according to the standard inventory identity (9), with depreciation at a fixed rate (\( \sigma \)) and additions due either to maintenance activity or expansion, capacity can be no smaller than a minimum plant size defined as a fraction of initial period capacity (10), and regional capacity is the sum of capacity at all mills (11).

Treatment of nonlinearities. Since the demand functions are linear, the willingness-to-pay integral in (1) is quadratic in mill receipts. This function also depends on current industry capacity (Capacity_{\text{EOR}}), which is endogenous and a further source of nonlinearity in the objective. To linearize the objective we employ two devices. Mill receipts (log demand) in each period are defined as a fraction of capacity in a set of discrete steps as (again with time subscripts suppressed):
The values employed in the initial iteration. Iterations yielded relatively stable values, depending below some tolerance. In application, we found that 5-7 values until the changes in capacity between iterations fall used, in turn, in the second iteration. This process continuing program involved approximately 120,000 constraints and investment) and was solved using a mixed integer linear programming algorithm. The model was coded in GAMS (Brooke and others 2003) and used the CPLEX solver. Typical base case problems without a public thinning program is a fixed vector of capacity utilization levels defined on the steps \([0, \mu, \mu+\xi, \mu+2\xi, \ldots, 1]\) with \(\mu\) defined above, CLEVELMILLS,STEPS is an endogenous (0,1) variable indicating which UTILSTEPS is employed, and CAPACITYMILLS is the current capacity of a mill from the previous \((i-1)\) solution iteration.

Since CLEVEL is a binary variable (0,1), the squared receipt term that would appear in the willingness-to-pay integral in the objective function can be expressed as a product of squared (exogenous) UTIL and CAPACITY terms and a single CLEVEL term, thus eliminating the potential nonlinearity in the receipt variables.

We eliminated the nonlinear product of receipt and capacity (in both the definition of receipt and in the objective function) by means of an iterative solution approach. In the first iteration, current capacity values in (13) and in the objective function \((\text{CAPACITY})^{(i-1)}\) are replaced by an estimate of equilibrium capacity. At the end of the first iteration, revised values of capacity are available from the solution [computed in constraints (8) and (11)]. These are used, in turn, in the second iteration. This process continues until the changes in capacity between iterations fall below some tolerance. In application, we found that 5-7 iterations yielded relatively stable values, depending on the values employed in the initial iteration.

The overall model formulation involves both discrete, binary variables (CLEVEL) and an array of continuous variables (including capacity, harvest, receipt, shipments and investment) and was solved using a mixed integer linear programming algorithm. The model was coded in GAMS (Brooke and others 2003) and used the CPLEX solver. Typical base case problems without a public thinning program involved approximately 120,000 constraints and 760,000 activities.

**ANALYSIS OF AN EXTENSIVE THINNING PROGRAM ON NATIONAL FORESTS**

Millions of acres of the national forests in eastern Oregon are at increased risk of damage by fire, insects and disease due to overly dense stocking, stemming in part from a long history of indiscriminate fire suppression and from cutting practices that focused on ponderosa pine. The “forest health” issue has attained some prominence in recent years as a result of several seasons of large wildfires and increased attention from the media and high-level policymakers. Thinning these stands, to accelerate growth of the residual stems, has been proposed as one possible (albeit controversial) approach to restoration. As Wilson, Maguire and Barbour (2002) observe, however, the economic feasibility of such a thinning program will depend on the costs of entering the affected stands and removing the material, the quality of the material to be removed and the impacts of large volumes of thinned material on market prices and the processing industry.

In this study we examine one hypothetical thinning program on national forests in eastern Oregon. The eastern Oregon forest sector model described above is run with and without the program and impacts assessed from differences. In the base market model (without thinning) public harvests are not sensitive to market conditions, entering in constraints (6) as exogenous additions to mill receipts. The thinning program is simulated by appending to the set of private inventory condition classes a group of sample plots representing national forest lands that might be thinned. Each public plot (expanded to represent its total area) has harvest and haul costs and associated sawtimber thinning volumes. Thinnings must be completed within the first 20 years of the projection (program duration). The model is then rerun and allowed to pick the plots and volume to be thinned over time.

Data for this analysis came from Wilson, Maguire and Barbour (2002). Plots from the Forest Service’s Continuous Vegetation Survey inventory system were examined to identify those that were “densely stocked”, defined as plots with stand density indexes (SDIs) in excess of 75% of the maximum SDI for their respective forest types. Tree removals were then computed for these stands to reduce SDI to 50% of the maximum, cutting from below (smallest first). In our market analysis we employ only the board foot volumes, measured for trees at least 9 inches in DBH, and use only the plots with non-zero board foot volume. Given the details of the thinning and plot characteristics, we computed harvesting costs (to the roadside) for the thinnings using the cost model of Fight, Gicqueau, and Hartsough (1999). In their study Wilson, Maguire and

---

7 Harvesting costs in the Fight, Gicqueau, and Hartsough (1999) model vary with a wide range of logging technology and site conditions. We use their default assumptions for all inputs except stem diameter, volume per acre to be removed, trees per acre to be removed, move-in distance, slope, and whether the harvest is a thin or not (all are thins in our case).

We assume that the Forest Service pays (subsidizes) all the costs of removing the non-merchantable volumes (live or dead trees less than 9 inches DBH) in these thinnings. Initially we also assume that sawtimber thinning volumes are made available with no stumpage fee. Thinning operators must pay the (substantially higher) harvesting costs of removing the thinning volumes but no other costs. We call this the “Full Subsidy” scenario. In a later simulation we add a simple fixed stumpage charge for the thinned sawtimber volumes to examine cost sensitivity.

Figure 2 shows total timber harvest in the base (no thin) and thinning (full subsidy) cases. Base case harvest continues to fall along historical trends until 2013. This is a reflection of the inability of both industry and NIPF lands to maintain recent historical harvest levels [see Adams and Latta 2003 for discussion]. Private harvests rise after the 2013 period but do not return to 1998-2002 levels for the remainder of the projection. Harvest by owner is detailed in figure 3. With a full subsidy thinning program, a large volume of public timber is harvested in the first program period (2001-2005). Initially public timber almost entirely replaces industrial harvest. Public harvest declines in subsequent periods, and the ownership mix moves back toward pre-program proportions. Private cut remains somewhat higher, however, as a result of the inventory “savings” allowed by the thinning program. The harvest trough in 2013 in the base case is displaced to 2043 in the thinning program and is less severe.

Changes in log consumption at mills mirrors patterns in the aggregate regional harvest, with higher volumes through the 2030 period in the thinning case. Driven by restrictions in private timber harvest, mill numbers in the base case decline rapidly after 2003, recovering to the 10-12 mill range after 2028 (figure 4). Under the thinning program, with higher harvest volumes, the current number of mills continues to operate until after 2033. Numbers then fall to the upper level of the base case range.

There appears to be no significant relationship between mill size and the volumes of thinnings taken from the national forests (bigger mills don’t necessarily get more of the thinnings). In the analysis, thinning use is governed instead by proximity to areas available for thinning (transport costs) and the harvest costs associated with the areas that are accessible to a specific mill.
Of the 1.5 million acres of densely-stocked stands included in the thinning program, we project that about 63% would be accessed for treatment under a full subsidy program (table 3). This would remove about 4.0 billion board feet of timber over a 20 year period, or about 91% of the total board foot volume on all areas in the thinning program. The total area in need of treatment (as identified by Wilson, Maguire and Barbour (2002)) less lands in wilderness and inventory roadless areas was about 1.95 million acres. Thus our analysis suggests that about 50% of this area could be treated under a full subsidy program.

In the thinning program simulation we assumed that the national forests made no charge for the thinned sawtimber and operators did not pay the costs of removing the non-merchantable material. Some procedure would be needed to select the firms to harvest the wood, but the only costs would be those of harvesting and transportation. A program of this sort should yield a relatively large area of thinned stands and a high volume of thinned sawtimber. We also examined a case in which additional charges were levied for the sawtimber. These additional costs might be viewed as fees for the value of merchantable logs removed or as a

---

**Table 3—Area (in acres) and volume (in MBF) treated and untreated in an example thinning program on national forests in eastern Oregon.**

<table>
<thead>
<tr>
<th></th>
<th>Acreage</th>
<th>MBF Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated in Period 1</td>
<td>332,552</td>
<td>2,269,835</td>
</tr>
<tr>
<td>Treated in Period 2</td>
<td>290,751</td>
<td>973,202</td>
</tr>
<tr>
<td>Treated in Period 3</td>
<td>112,962</td>
<td>269,934</td>
</tr>
<tr>
<td>Treated in Period 4</td>
<td>235,370</td>
<td>455,050</td>
</tr>
<tr>
<td>Total Treated</td>
<td>971,635</td>
<td>3,968,020</td>
</tr>
<tr>
<td>Percent of Thinning Program</td>
<td>63%</td>
<td>91%</td>
</tr>
<tr>
<td>Total Untreated</td>
<td>564,180</td>
<td>412,995</td>
</tr>
<tr>
<td>Percent of Thinning Program</td>
<td>37%</td>
<td>9%</td>
</tr>
<tr>
<td>Total in Thinning Program</td>
<td>1,535,815</td>
<td>4,381,015</td>
</tr>
<tr>
<td>Total Overstocked Area</td>
<td>1,953,149</td>
<td></td>
</tr>
<tr>
<td>Percent overstocked in thinning program&lt;sup&gt;a&lt;/sup&gt;</td>
<td>79%</td>
<td></td>
</tr>
</tbody>
</table>

MBF = thousand board feet

<sup>a</sup> Area involved in the thinning program included only those acres with non-zero board foot volume
proxy for the cost of removing unmerchantable trees and dead wood as part of the thinning activity (if these costs were proportional to the sawtimber volumes removed). Figure 5 shows the public harvest results from the imposition of a $150 per MBF stumpage charge. The thinning program essentially represents a stock that can be drawn down at any time during the 20 year program period. Increasing the cost of this material causes thinning harvests to be partly postponed into the third and fourth periods.

DISCUSSION

Given the decline in private harvests in eastern Oregon projected in the base case, an extensive thinning program on national forests in eastern Oregon over the next two to three decades would be well-timed. Results suggest that about 2/3 of the overstocked area that has trees large enough to yield board foot volume would be thinned under a full subsidy program. Total harvest in the region would eventually fall to levels foreseen in the base case, but not until after 2030. The program and its aftermath would also keep mill numbers at current levels for an additional 30 years. The analysis indicated that public cut from the thinning program would largely replace harvests from industrial lands and a sizable portion of the cut from NIPF lands in the first program period. Thereafter, private harvests would move back toward base levels but with a larger inventory and somewhat higher long-term harvest potential. Geographically the largest portion of untreated/unthinned plots and those with no board foot volume are located in the central Oregon (eastern Cascades subregion). Because only those plots with positive board foot thinning volume were included in the analysis, we obtained a higher proportion of area treated than is commonly believed possible for such a program.

To develop these projections we constructed a model of the eastern Oregon forest products sector at the log market level with explicit representation of individual mills. While this model appears to hold some promise for examining the increasingly complex questions related to forest health restoration measures on public lands in eastern Oregon, we recognize that it has limitations that will be challenges for future research. Unlike many past studies, the log demand relations were estimated from a profit function model that included capacity explicitly as a measure of quasi-fixed capital stock. The estimation did not, however, provide values for the critical parameters for minimum operating rate and minimum mill size. Also, absent cost and operating data at the mill level, we estimated mill-level demand by proportioning the regional log demand curve, assuming homogenous logs and uniform products produced by all mills. Ideally, these estimates would be derived for individual mills recognizing efficiency, cost and output quality differences across the mill population.

LITERATURE CITED


THE POTENTIAL OF ARCLOGISTICS 3 ROUTE FOR SOLVING FORESTRY TRANSPORTATION PROBLEMS

Hamish Marshall and Kevin Boston

ABSTRACT

For many years the forestry industry has utilized the data storage, analytical and data visualization power of geographical information systems (GIS) to improve their operations. One area where the forest industry has been slow to utilize the power of GIS is in the planning and management of their transportation systems. As the industry moves more towards customer driven supply chain management systems, the forest industry will need to intensify the management of their transportation systems. This paper reviews ESRI’s ArcLogistics software suitability to solve modern log truck routing and scheduling problems faced by the forest industry. The results of the scenarios show that the functionality of ArcLogistics is probably insufficient to be used by large forestry companies for their day-to-day scheduling. However, it may prove to be a cost effective tool for medium to small forest companies to evaluate the effects of transport and inventory policies.

INTRODUCTION

The transportation of logs from the forest to the mill is a main component of the forest supply chain. Log transportation is the single largest cost for many forest companies around the world. In the southern United States, log transportation accounts for nearly 50% of the delivered cost of wood fiber (McDonald, Rummer, Taylor and Valenzuela, 2001). In New Zealand, forest industry log transportation accounts for 20 to 30% of the seedling to mill-door discounted costs (Carson 1989).

The size of logging trucks makes them an imposing vehicle on the public roads; the result is a negative public interaction with the forest products industry. Accidents involving logging trucks appear to draw considerable media attention regarding the dangers of logging trucks (Greene and Jackson 1995). There is also an increased public intolerance to the sight of logging trucks within residential areas. The forestry industry needs to develop new transportation plans that avoid these residential areas to reduce potential negative interactions.

Murphy (2003) solved the transportation plan for two case studies in New Zealand that showed similar reductions to the other two modelling efforts with a reduction in fleet size between 25 and 50%.

Despite the potential gains that can be made through improved transportation planning and the considerable interest by forest industries worldwide (Barrett, Prisley and Shaffer 2001, Cossens 1993, Palmgren 2001) there are very few logistic planning systems available to the forestry industry. Probably the most well known of the commercially available systems is ASICAM. It produces a daily plan using a heuristic-based simulation model that assigns and builds routes to the available trucks and loaders to satisfy supply and demand constraints. ASICAM resulted in a 32% reduction in truck fleet size in Chile (Epstein, Morales, Sera, and Weintrab 1999).

Additional work has been completed in Finland and Sweden with the development of EPO and RUTT-OPT, which are customized forest logistics planning systems (Palmgren 2001). Two case studies were completed using
RUTT-OPT that showed a large forest company could reduce its fleet size by 30% through improved truck utilization. A small wood chip company was able to reduce its total distance driven by 8%, achieving both the cost savings and improving public safety with a reduction in road traffic (Palmgren 2001).

There are a number of other transportation software systems commercially available that are not specifically designed for solving forest logistic problems, these include TransCAD developed by Caliper Corporation, SHORTREC developed by ORTEC and ArcLogistics 3 Route developed by Environmental Systems Research Institute (ESRI), which is reviewed in this paper.

ESRI is the world leader in Geographical Information Systems (GIS) and holds a 72% of the market share for GIS products (Ecity, 2003). ArcLogistics 3 Route (ArcLogistics) is a stand-alone product developed to solve transportation and routing problems for delivery businesses where a truckload of goods is picked up at one location and then delivered to a number of customers at different locations. ArcLogistics Route improves these businesses by delivering goods or services more efficiently with reduced costs and improved customer service. Unlike ESRI’s other products, ArcLogistics does not require a full ArcGIS licence to run. ArcLogistics uses the Dynamap/Transportation U.S street data which is included with the package. The database includes node elevations and features such as one-way streets, physical turn restrictions, calculated speed information and a geo-reference database that allows the precise location of the customer to be found using the customer’s address.

ArcLogistics requires the user to describe the vehicle type and driver. This includes information on individual truck capacities, cost information in terms of fixed cost, $ per mile, $ per hour and $ per overtime hour. Users can also specify working hours, details about lunch breaks, as well as information on start and end locations and times. This data is used to determine optimal routes.

Orders can be entered using the user interface or by importing them from a database. Once all the necessary information has been entered, the optimal routes and schedule are generated. ArcLogistics develops an origin-to-destination matrix that creates the possible routes from the supply to the demand nodes. Good solutions are found to the problem using a tabu search heuristic. Although the quality of solutions produced by ArcLogistic were not tested in this paper, the tabu search heuristic has been found to produce excellent solutions to both transportation problems (Gendreau, Hertz and Laporte 1994) and other forestry problems (Bettinger, Graetz, Boston, Sessions, and Chung 2002). The results are displayed using both graphical and tabular reports. The tabular reports were found to be the most useful method for reporting for the analysis completed in this paper. If one were using ArcLogistics for day-to-day scheduling and routing of a fleet of trucks, the quantity of information displayed in the graphical reports would be extremely useful.

ArcLogistics was developed to solve problems where the truck has a single pickup point and multiple delivery points per load. These differ from the logistic problems that normally exist in the forestry industry where either a truck is filled at a single location and delivered to another location, or the more complex case (as in Scandinavia) where the order is collected from a number of spatial locations and delivered to one location. The objective of this paper was to investigate the potential of ESRI ArcLogistics 3 Route software for both operational scheduling and routing of trucks in the forest industry.

METHOD

To test the suitability of ArcLogistics to solve simple forestry transportation problems, a set of small hypothetical test problems were developed. ArcLogistics will need the following functionality to successfully solve these problems:

- Be able to route trucks from forest to mill and back to the forest.
- Restrict available routes from the network.
- Restrict loading and unloading hours to meet local restrictions and mill operating hours.
- Specifying the maximum driver hours.
- Place different priorities on customer orders, so that high priority customers are served first if there is a deficit of trucks.

The hypothetical problems used the Oregon State University McDonald-Dunn Research Forest located near Corvallis, Oregon. The model has ten harvesting sites (supply nodes) and five mill locations (demand nodes). These include four local sawmills and one pulp mill. The model also had one garaging location where the trucks had to return at the end of the day. For this problem, a total of 60 orders were generated. Each order specified the supply location, the address of the customer and the total volume/weight of the order. The “base model” specified the following constraints:

- Trucks can only be loaded between 3 am and 4 pm in the forest.
Trucks can only be unloaded between 5 am and 6 pm at the mills.

• Trucks must return to their garages at the end of the day.
• The drivers can only drive for 12 hours a day.
• The capacity of the trucks is 40 tonnes. No split loads (loads to two destinations) were allowed.

Three scenarios were derived from this base model to test different features of ArcLogistics. Scenario one was designed to test ArcLogistics functionality with a surplus and deficit supply of trucks. This scenario was first run with three fewer trucks available than in the base model. It was then rerun with 5 more trucks then in the base model to simulate a situation where there was an oversupply of trucks.

Scenario two was developed to test ArcLogistics functionality regarding a mill’s desire to operate its log inventory using a just-in-time (JIT) inventory control system. Logs are required to be delivered within a set time period. The orders for one of the mills in the base model were altered so that each order had to be delivered within a predefined one-hour time slot. These delivery times were evenly distributed through the day. This scenario was designed to simulate having orders to be delivered to a mill that was operating a JIT inventory management system.

Scenario three was designed to test the priority orders system in ArcLogistics. The model used an insufficient supply of trucks to complete all loads but a subset of the orders had a higher priority for fulfillment than the others. This will allow for the company to satisfy customers based on a priority system.

RESULTS

The scenarios were compared using several measures produced from the ArcLogistics reports. The first measure was simply the total cost of the deliveries made. The second was the percentage of the total orders that were actually delivered. The two other measures were used to measure truck utilization; average mileage traveled per truck, average proportion of the day (12 hours) spent driving per truck. Table 1 gives a summary of these four measures for each of the scenarios.

**Base scenario**

The base model took 50 seconds to solve using the default setting. The model produced a solution in which all the orders were allocated to trucks while not breaking any of the constraints. The total cost for the base model was $3059.79. The utilization average for the trucks, the amount of time the truck was driving, was 82% of the 12-hour day.

**Scenario 1**

ArcLogistics performed as would be hoped, when a surplus of trucks existed. ArcLogistics allocated the orders to trucks in a manner that minimizes the trucks required. This is opposed to allocating the orders to all the trucks, creating a situation in which the majority of the trucks are only utilized for part of the day, which is in most cases an undesirable situation. When too few trucks are included in the model, ArcLogistics chose to not deliver the orders that have the highest delivery cost, those orders that took the longest time to deliver. The total cost when there was a deficit of trucks was just over a $1000 less than the base scenario. This is due to the lack of trucks, meaning that not all orders were delivered and hence no transportation cost was incurred from the undelivered orders. The average cost of delivery for an order also went down from $41.91 for the base scenario to $32.49 when only 5 trucks were supplied to the model. This indicates that the model is, as stated in the documentation, minimizing cost.

**Scenario 2**

The results from the JIT application showed that this management option usually requires more resources as it contains additional delivery time constraints. Upon detailed analysis of each truck’s delivery schedule, it was found that

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Number of Trucks Used</th>
<th>Total Cost</th>
<th>Percentage of Orders Delivered</th>
<th>Average Miles Traveled per Truck</th>
<th>Average Drive Time per Truck</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>8</td>
<td>$3059.79</td>
<td>100 %</td>
<td>241</td>
<td>82 %</td>
</tr>
<tr>
<td>1a</td>
<td>5</td>
<td>$1917.19</td>
<td>81 %</td>
<td>345</td>
<td>84 %</td>
</tr>
<tr>
<td>1b</td>
<td>8</td>
<td>$3059.79</td>
<td>100 %</td>
<td>241</td>
<td>82 %</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>8</td>
<td>$3076.59</td>
<td>100 %</td>
<td>243</td>
<td>82 %</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>$2445.01</td>
<td>85 %</td>
<td>220</td>
<td>87 %</td>
</tr>
</tbody>
</table>

Table 1—Results from the Five Different Scenarios
ArcLogistics had successfully complied with the “just-in-time” constraints placed on the individual orders. The total cost for delivering the order had increased to $3079.59. Under this scenario, the average number of miles traveled per truck increased, but the average utilization percentage of the 12 hours spent driving remained the same. The scenario did show that ArcLogistics could be used to investigate the effects that JIT management at sawmills will have on transportation costs and logistics.

Scenario 3
In Scenario 1, when the model had too few trucks to deliver all the orders, the model did not satisfy the orders with the longer distances. However in Scenario 3 these orders had a higher priority placed on them than the other orders. ArcLogistics delivers these orders first and did not fulfill the low priority orders. This allows ArcLogistics to develop transportation plans that will meet those critical orders such as wood needed to meet an export order to avoid demurrage or deliver wood to a wood processing facility that must have raw material to maintain their operations. The total cost of this scenario was less than the base model for the same reason as stated in scenario 1. In this scenario the average cost per order was $39.43. The truck utilization remained 87% of the 12 possible hours that they could drive per day.

DISCUSSION AND CONCLUSION
The scenarios developed and solved in this paper show that although ArcLogistics was not designed to solve forestry transportation problems, it can be adapted to solve many of the common problems, but it does have some limitations that may reduce its use in the forestry industry. This paper does not include an examination of solution quality of the tabu search heuristic.

The first major limitation is that it uses the Dynamap/Transportation U.S street data that only includes public roads. For most forest products companies, the majority of the roads in their transportation networks are privately owned and will not be included in the Dynamap/Transportation U.S street data. Although it is possible to add roads to this database, the process seems to involve some advanced GIS skills and expenses of further data capture. In this case study the location of the landings in the forest had to be placed on public roads, which were included in the Dynamap/Transportation US street data. This is far from the reality that actually exists in the McDonald Dunn Research Forest, where many of the harvest units can only be reached using forest roads that are not included in the ArcLogistics street dataset. Another limitation of ArcLogistics is the inability to customize the formulation of the model. The model can only minimize cost and does not allow users to define alternative objective functions containing utilization, safety and other economic variables.

Despite these limitations, the results of this research have shown that ArcLogistics could be used in the forest industry. Given the current functionality of ArcLogistics, it is hard to imagine a large forest company implementing it for their day-to-day operational scheduling and routing. However, the cost of $US 12,000 (as of December 2003) makes ArcLogistics a very cost competitive tool for medium to small size forestry companies to plan and evaluate the cost and efficiency of different transportation and inventory management policies when customized versions can cost millions of dollars to develop and implement.

ESRI indicates that the new version of ArcLogistics, which is due to be released in a beta version sometime next year, will come in the form of an ArcGIS extension such as Spatial Analyst. This will mean that using a private road network database will be significantly easier. It will also potentially allow for more customization of objective functions, constraints and the optimization models. This means that the new version of ArcLogistics may eliminate many of the limitations of the current version of ArcLogistic that are present when trying to solve forestry transportation problems.

LITERATURE CITED


This page is intentionally left blank.
USING SPATIAL INDICES TO ESTIMATE WOOD PROCUREMENT COSTS

Luc LeBel, Sylvain Théberge, Kim Lowell, and Vincent McCullough

ABSTRACT

In many forest regions, cutting blocks are increasingly being dispersed over large procurement areas. Basic operational analyses make it clear that spreading operations over a larger area increases costs. However, predicting operating costs is time-consuming because a precise harvesting plan has to be laid out to measure forest logistics parameters such as total road distance and cut block size. In an attempt to make the process of evaluating different management scenarios more efficient, a GIS-based method of estimating the costs of forest operations was developed. This paper considers the problems and possibilities regarding Spatial Indices (SI), and attempts to demonstrate that 1) forest mosaics can be decomposed into spatial characteristics that are quantifiable by means of SIs, 2) a combination of SIs, used at optimal conditions enables analysts to define quantitatively the spatial characteristics of a forest mosaic, and 3) relationships can be found among SIs and wood procurement parameters. The procurement model uses cost functions based on the relationship between spatial indices that quantify factors such as cut-block dispersion and landscape heterogeneity, and operational factors such as skidding and hauling distance that were developed through multiple regression analysis.

INTRODUCTION

Forest resources managers must now adapt forest practices to reflect spatial patterns from natural disturbances and natural biophysics environment (Hessburg and others 1999, Harvey and others 2002). They are also required to distribute forest interventions in space and time to accommodate different forest users. The “mosaic” management system constitutes a forestry approach that makes this possible (Quebec 2003).

The mosaic management system consists of a forestry approach that makes possible the distribution of forest interventions in space and time so that they reflect patterns from natural disturbances and natural biophysics environment. The dispersion of logged areas is increasingly required by forest users. Indeed, it is recognized that a high concentration of cut blocks does not always support the preservation of diversified quality landscapes or favorable conditions for multiple land uses. The mosaic management method is expected to maintain a diversity of favorable wildlife habitats and minimize the impact of forest operations on the landscape’s aesthetic quality. It will also contribute to long-term biodiversity preservation (Quebec 2003).

On the other hand, for the companies carrying out forest operations, this type of management has, in the short term, a negative impact on their profitability. Indeed, it has already been shown that, in the short run, a substantial increase in costs can be observed. These costs are, for the most part, the consequence of a higher investment in capital and an increase in the costs of maintaining the road network (O’Hara and others 1989, Zundel 1992, Gingras 1997, Nadeau and others 2002). As a consequence of the uncertainty surrounding the breadth of additional costs associated
with this management method and the difficulties of defining the operational details with precision, the mosaic management system has been used with caution by the forest industry.

Each territory is characterized by a mosaic that is unique. In order to evaluate management strategies and compare future forest states, the characterization of the initial spatial structure of the forest is necessary (Baskent 1998). Moreover, a forest land manager should be able to characterize the initial (“natural”) mosaic in order to overlay an operational mosaic that reflects local objectives: forest yield, timber supply harmonization and security, wildlife habitat quality, other forest users, and timber supply costs. From an operational point of view, the characterizing approach must correspond to a rapid and simple analytical method that allows for binding the characteristics of intervention mosaics with certain parameters that influence the timber supply cost. The operational parameters to consider are the harvesting and skidding costs, transportation cost, and road network construction and maintenance costs. All of these parameters can be related to certain spatial variables such as shape, size and dispersion of cut-blocks.

In land management, the use of spatial analysis tools such as spatial indices (SIs) should facilitate the application of new spatial and time standards for forest operations and minimize the management of land with human perception (D’Eon and Glenn 2000). For example, rather than defining (qualitatively) a mosaic as fragmented with small, simply shaped polygons, one could say (quantitatively) that the mosaic has a fragmentation of 85.5%, a shape complexity of 27.0, and that the polygons have an average size of 12.6 ha.

The field of Landscape Ecology has described many SIs that allow for the quantification of landscape structure (Mladenoff and others 1993, Gustafson 1998, Tischendorf 2001) and can be used to characterize forest mosaics. The science of Landscape Ecology examines, among other things, the development and the dynamics of spatial heterogeneity, spatial heterogeneity management, and the influences of spatial heterogeneity on biotic and abiotic processes (Domont and Leduc 1994).

The definition of a landscape in the discipline of Landscape Ecology can be very diversified but, simply put, a landscape can be regarded as a spatially heterogeneous area (Turner 1989). The description of a forest mosaic corresponds to this definition. Effectively, a forest landscape consists of a mosaic of various stands that are homogeneous relatively to their dendrometric and ecological characteristics.

The forest mosaic can be characterized according to the variation of size and shape, spatial distribution, density, and stand characteristics (species, age, height, crown closure) (Ripple and others 1991, Mladenoff and others 1993, Baskent and Jordan 1995, Kleinn 2000). SIs take into account individual patches (size, number, form or density) or their vicinity. They study the landscape’s (or map’s) composition and spatial pattern (Gustafson 1998). As for composition, they examine the number of classes on the map and/or the proportion and the diversity of certain classes according to the whole landscape. Alternatively, the spatial pattern is analyzed according to the average, the median, the variance or the frequency distribution of a patch’s surface or linear measurements, their juxtapositions (He and others 1999). SIs can be categorized according to the following categories (this list is not exhaustive but shows the main uses for SIs):

- **Patch Size:** The size is the simplest measurement of the spatial character of a patch. The majority of SIs incorporate directly or indirectly the patch size (McGarigal and Marks 1995, Gustafson 1998).

- **Patch Shape Complexity:** Shape complexity refers to the geometry of a patch; this can be simple and compact or irregular and large. Shape is a spatial attribute that is very difficult to define with SIs because there is an infinite number of patch shapes possible. This spatial aspect is usually summarized or generalized by a shape complexity index taking into account the complexity of a whole mosaic. The measurement of shape complexity is usually based on the ratio perimeter-surface relationships between landcover proportion and indices of landscape spatial pattern (Forman and Godron 1986, Gustafson and Parker 1992, Baskent and Jordan 1995, McGarigal and Marks 1995, Gustafson 1998).

- **Core Area:** As indicated by its name, Core Area represents the inside area of a patch, after a buffer zone is specified by the analyst; it is the patch area not affected by the edge. This edge distance is defined by the investigator according to the organism or phenomenon observed and it can be fixed or variable, according to the edge types. Core Area integrates, in a single measurement, the size, the shape and the edge effect distance (Ripple and others 1991, Baskent and Jordan 1995, McGarigal and Marks 1995, Gustafson 1998, Potvin 1998).

- **Isolation/Proximity:** Isolation/proximity refers to the tendency of patches to be relatively isolated from other
patches of same or similar classes. Isolation can be calculated in terms of size and proximity to the neighboring patches, and this for each patch in the mosaic (Gustafson and Parker 1992, McGarigal and Marks 1995, Gustafson 1998).

- **Contrast**: Contrast refers to the relative difference between adjacent patch types. For example, mature forest, adjacent to immature forest will have an edge contrast that is weaker than mature forest adjacent to a field in culture. This type of SI can be calculated by assigning a contrast weight to each type (McGarigal and Marks 1995, Gustafson 1998).

- **Dispersion**: Dispersion refers to the tendency for the patches to be regularly dispersed, non-random (D’Eon and Glenn 2000) or aggregated. This index is often calculated as a function of the variability of nearest neighbor distance measures. A small standard deviation relative to the mean implies a fairly uniform or regular distribution of patches across landscapes, whereas a large standard deviation relative to the mean implies a more irregular or uneven distribution of patches (McGarigal and Marks 1995, Baskent 1998, Trani and Giles 1999).

- **Contagion/Interspersion**: The contagion refers to the tendency of patches of the same type to be spatially aggregated. Contagion does not take the patch into account as such, but rather the extent to which cells of similar classes are aggregated. There are several ways of calculating these indices but in general, an algorithm determines the probability of finding cells or adjacent patches of the same types (McGarigal and Marks 1995, Gustafson 1998, Hargis and others 1998, He and others 1999). The lacunarity is also a measurement used for the calculation of contagion and interspersion. It is a measurement borrowed from fractal geometry which calculates landscape contagion at several scales (McGarigal and Marks 1995, Gustafson 1998).

- **Subdivision**: Subdivision refers to the scattering of patches. This SI category does not refer to the size, form, relative position or spatial arrangement of patches as such. However, since the subdivision generally affects patch spatial organization, it is difficult to isolate subdivision as an independent measurement. This index can be evaluated according to the number or density of patches, average patch size, aggregation or patch size distribution (McGarigal and Marks 1995, Gustafson 1998, Jeager 2000). When applied to the class level, these SIs can be used to determine the fragmentation of a patch type. Moreover, applied to a whole landscape, they can be used to quantify landscape texture (granulation). A landscape with a fine texture is characterized by its many small patches while a landscape with a coarse texture shows some large patches (McGarigal and Marks 1995, Trani and Giles 1999). It can be based on strict adjacency, a threshold of distance or a specific function of the distance between patches. Several SIs can be derived from patch connections (McGarigal and Marks 1995).

Since the beginning of the 1980s, several SIs were introduced for various uses (Turner 1989, McGarigal and Marks 1995, Jeager 2000). In spite of a significant interest in the use of SIs, there is still some uncertainty as to their effectiveness for quantifying the characteristics of “real” landscapes. Indeed, SIs have mainly been examined in relation to simulated (computer generated) landscapes and their behavior on real landscapes is not well understood. Therefore, the behavior and sensitivity of SIs relative to different spatial patterns of “real” landscapes are obscure. The knowledge of SIs’ behavior and sensitivity in relation to different area sizes of “real” landscapes is also sparse. Furthermore, while characterizing a landscape, particular care must be taken in regard to scale, resolution and landscape extent (McGarigal and Marks 1995).

Since each SI has a theoretical extent and a frequency distribution that has been analyzed using simulated landscapes (Gustafson 1998, Hargis and others 1998, He and others 1999), it is difficult to determine which SIs are the most valuable on real landscapes. The empirical analysis of “natural” variations of SIs remains one of the greatest challenges confronting landscape pattern analysis (McGarigal and Marks 1995).

The lack of knowledge of SIs and the difficulties associated with the interpretation of their responses to spatial pattern is a significant obstacle to the goal of defining landscapes objectively and using SIs for other forestry applications such as evaluating wood procurement costs. Indeed, these problems have caused an incomplete integration of the principles of landscape ecology in resources management (Davidson 1998, Gustafson 1998, Tischendorf 2001, Plante 2002).

If several SIs are correlated, show interactions among themselves, and measure multiple components of spatial patterns, analysis is difficult (Hargis and others 1998, Tischendorf 2001). Consequently, consideration of several SIs measurements, individual and specific to one heterogeneity variable, can be very instructive (Li and Reynolds 1995). Thus, the use of SIs is effective for the comparison of different landscapes, the same landscape at different scales (McGarigal and Marks 1995, Trani and Giles 1999). It can be based on strict adjacency, a threshold of distance or a specific function of the distance between patches. Several SIs can be derived from patch connections (McGarigal and Marks 1995).
times or the same landscape under different management strategies (Gustafson 1998).

Considering the problems and possibilities regarding SIs, this study will demonstrate that 1) forest mosaics can be decomposed into spatial characteristics that are quantifiable by means of SIs, 2) using a combination of SIs at optimal conditions should enable us to define quantitatively the spatial characteristics of a forest mosaic, and 3) relationships can be found among SIs and wood procurement parameters.

METHODS

Study Area and Data

To determine the behavior and the sensitivity of SIs on real forest mosaics, a sample of forest cover maps with different spatial configurations was needed. Knowing that spatial patterns of forest mosaics vary from one geographical area to another, we asked for, and received, the collaboration of three forest companies: Abitibi-Consolidated, Bowater and Smurfit-Stone. These companies’ lands are located in distinct areas of Quebec’s (Canada) public forest. By each providing two maps that they judged different in respect to spatial configuration of forest types (for example, patch diversity, patch geometry, patch size, etc.), these companies’ representatives provided six forest mosaics that had different structures. These differences were the consequence of factors such as topography, fire regime, latitude, or disturbances caused by spruce budworm (*Choristoneura fumiferana* (Clem.)). In addition, the companies provided digital map layers containing historical dispersion patterns of forest operation layouts: roads, harvested blocks, wood yards, camp locations, etc. These map layers provided essential information to establish relations between forest mosaic spatial characteristics and wood procurement costs.

These datasets, used for forest management, include a geometrical and descriptive database with common elements. The maps are based on the geodetic reference NAD 83, cartographic projection MTM and a scale of 1/20 000.

Data

The forest cover maps described previously are vectorial categorical cartographic documents. They consist of discrete polygons that represent homogeneous forest vegetation zones (vegetation types). In order to convert the map from vectorial format to raster format, the software *ArcView* 3.2 (*ArcView GIS 3.2* is a software produced by the company Environmental Systems Research Institute, Inc. Copyright ©1995. 2003 ESRI) was used. The images for one of the three companies are presented in Figure 1. The raster format of 10m X 10m resolution used to perform the analysis does not significantly affect the precision of the forest cover map, however it does simplify SI programming.

Quantitative Characterization of Forest Mosaics

Window Size for the Calculation of SIs—The spatial pattern detected in an ecological mosaic is related to the scale (Forman and Godron 1986, Turner 1989, McGarigal and Marks 1995); the ecological concept of scale includes...
extent and grain. The extent and grain define a resolution’s higher and lower limits and any inference relating to the dependence of a system on a given scale is driven by the extent and grain. Consequently, it is essential that the extent and grain represent accurately the ecological phenomenon or organism under investigation. Otherwise, the landscape patterns detected will have little significance and will probably result in erroneous conclusions. Unfortunately, this ideal manner of defining scale is not very realistic. In practice, extent and grain size are often dictated by the scale of the image used (for example, the resolution of remotely sensed image) or data-processing capabilities. For this study, the SIs were tested empirically using different scales. This was done by keeping the grain (pixel) size constant and applying each SI in a moving window; an area unit defining a specific landscape extent and moving in regular intervals across the forest mosaics.

Choice and Calculation of the Selected SIs—Ten SIs were selected to quantify different spatial aspects considered to be essential for the characterization of various landscapes (Table 1). Patch Richness Density (PRD), which tallies the classes present in the landscape, and Patch Density (PD), which counts the number of patches, were used to quantify landscape composition. Regarding spatial patterns, Edge Density (ED), Landscape Shape Index (LSI), and Area-Weighted Mean Shape Index (AWMSI) provide a measurement of patch complexity in the landscape. Contagion Index (CONTAG) gives a quantitative evaluation of landscape fragmentation while Patch Cohesion Index (COHESION) measures connectivity among patches of the same class in the landscape. Regarding patch area size, Mean Patch Area Distribution (AREA_MN) indicates the average patch size in the landscape whereas Mean Core Area Distribution (CORE_MN) provides an average of the patch core area for the patches in the landscape. Finally, the Range of Euclidean Nearest Neighbor Distance (ENN_RA) quantifies the isolation/proximity of patches in the landscape, i.e., the tendency for the patches to be more or less isolated from patches of the same class. This collection of SIs shows that the forest mosaic can be broken up into characteristics that can be quantified to define its spatial patterns and composition. McGarigal and Marks (1995) detail the mathematical expressions for these and other SIs.

Each of the SIs documented in this section can be used with a vectorial or raster data format. However, since the SI’s calculation is performed on raster images, the explanation and comments on SIs and their mathematical expressions (Table 1) will be done for this format of data.

Frequency Distributions and Statistics—For each of the six cartographic layers, the ten selected SIs were calculated within a moving window of 50 ha, 200 ha, and 400 ha. The SIs were calculated only for windows where each patch had an attribute and the window did not overlap the map’s external limits. From the SIs results, the frequency distributions were generated. From these distributions, the basic descriptive statistics were determined.

To understand SIs’ behavior and determine if they can detect visual spatial configuration differences between different landscapes, one can test if two independent empirical distributions come from two different populations. To determine this and to verify the validity of each SI, the Kolmogorov-Smirnov two-sample test (K-S) was used. The K-S test was conducted using the software S-Plus. (S-Plus 6.0 is a software produced by the company Insightful Corporation. ©2002 Insightful Corporation.) The threshold for the rejection of the hypothesis that two distributions are identical was 99%.

Use of SIs to Predict Key Wood Procurement Parameters

Using detailed operational maps of each of the procurement areas studied, a spatial analysis was conducted. The spatial analysis of the past operation layouts helped establish which factors - road construction, harvesting, transportation, equipment displacement (moving) and planning - contributed to wood procurement cost. It also permitted determining at what scale - whole landscape, operational area, harvesting sectors, or cut blocks - the SIs should be used for effective cost quantification. Moreover, the analysis highlighted which spatial features - roads, harvesting sectors dispersion, cut block shapes, etc. - are significant for the characterization of the operational mosaic and which SIs could contribute to its quantification.

Since the cost of forest operations is usually based on the specific forest class of interest, for example, the cut blocks for the cost of harvesting, or the distance between harvesting sectors and the mill for the cost of transportation, a binary model was used to characterize the operational mosaic with SIs.

This analysis allowed the identification of SIs that are most suitable for calculation of different forest operation features costs. These are Contagion Index, Euclidean Nearest Neighbor, Edge Density, Class Area (Patch Area, but for a given class), and Patch Density. Once the SIs had been selected to quantify different spatial aspects of operational features, the multiple regression analysis technique
Table 1—Summary of the selected landscape metrics used for the analysis.

<table>
<thead>
<tr>
<th>Name</th>
<th>Characteristics</th>
<th>Mathematic expression</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patch Richness Density (PRD)</td>
<td>Measures the number of patch classes in the landscape</td>
<td>( PRD = \frac{m}{A} (10,000)(100) )</td>
<td>Number / 100 ha</td>
</tr>
<tr>
<td>Patch Density (PD)</td>
<td>Measures the number of patches in the landscape</td>
<td>( PD = \frac{N}{A} (10,000)(100) )</td>
<td>Number / 100 ha</td>
</tr>
<tr>
<td>Edge Density (ED)</td>
<td>Measures the total length of edge segments in the landscape</td>
<td>( ED = \frac{E}{A} (10,000) )</td>
<td>m/ha</td>
</tr>
<tr>
<td>Landscape Shape Index (LSI)</td>
<td>Measures the total length of edge segments in the landscape and compares it to a square standard</td>
<td>( LSI = \frac{E}{\sqrt{A}} )</td>
<td>Without units</td>
</tr>
<tr>
<td>Area-Weighted Mean Shape Index (AWMSI)</td>
<td>Measures the average patch shape (in the landscape) by weighting the patch area so that larger patches have larger weight than smaller patches</td>
<td>( AWMSI = \sum_{i=1}^{m} \left( \frac{P_i}{\sqrt{a_i}} \right) \left( \frac{a_i}{A} \right) )</td>
<td>Without units</td>
</tr>
<tr>
<td>Contagion Index (CONTAG)</td>
<td>Measures the relative aggregation of the landscape</td>
<td>( \text{CONTAG} = \left( \frac{\sum_{i=1}^{m} \sum_{j=1}^{m} \left( P_i g_{ij} \right) \ln \left( P_i \right) g_{ij}^{m \sum_{j=1}^{m} g_{ij}}}{2 \ln (m)} \right)^{100} )</td>
<td>Percentage (%)</td>
</tr>
<tr>
<td>Patch Cohesion Index (COHESION)</td>
<td>Measures the connectivity of patches of the same class in the landscape</td>
<td>( \text{COHESION} = \left( \frac{1}{\sum_{i=1}^{m} P_i^{<em>} \sqrt{a_i^{</em>}}} \right) \left[ 1 - \frac{1}{\sqrt{Z}} \right]^{-1} \times 100 )</td>
<td>Percentage (%)</td>
</tr>
<tr>
<td>Mean of patch Area Distribution (AREA_MN)</td>
<td>Measures the average patch area size for the landscape</td>
<td>( \text{AREA}<em>M_N = \frac{\sum</em>{i=1}^{n} \left( \frac{a_i}{10,000} \right)}{n} )</td>
<td>Ha</td>
</tr>
<tr>
<td>Range of Euclidean Nearest Neighbor Distance (ENN_RA)</td>
<td>Measures the range of the distance between patches of the same class in the landscape</td>
<td>( \text{ENN}<em>R_A = h</em>{i_{\max}} - h_{i_{\min}} )</td>
<td>Meter</td>
</tr>
<tr>
<td>Mean of Core Area Distribution (CORE_MN)</td>
<td>Measures the average patch area size for the landscape, excluding an edge buffer (specified by the user)</td>
<td>( \text{CORE}<em>M_N = \frac{\sum</em>{i=1}^{n} a_i^{c} \left( \frac{1}{10,000} \right)}{n} )</td>
<td>Ha</td>
</tr>
</tbody>
</table>

\( m \): number of patch types in the landscape  
\( A \): landscape total area (m\(^2\))  
\( E \): landscape total length of edge segments  
\( N \): number of patches in the landscape  
\( p_i \): patch perimeter (m)  
\( a_i \): patch area (m\(^2\))  
\( \% \): landscape proportion occupied by patch type i  
\( g_{ij} \): number of adjacency between pixels of type i and j  
\( P_i^{*} \): patch perimeter (in pixel)  
\( a_i^{*} \): patch area (in pixel)  
\( Z \): total number of pixel in the landscape  
\( h_i \): shortest Euclidean distance from the edge of a patch and a patch of the same class  
\( a_i^{c} \): patch core area (Note: a edge buffer is specified by the user, in this study: 50 meters)
was used to determine the correlation between the SIs and the measured operational features. Also, the multiple regression technique permitted the generation of functions using SIs to calculate the operational features. The result is a predictive model. Detailed historical data provided by the companies allowed for the comparison between actual and predicted values. Finally, the cost functions developed were constructed and tested for robustness and sensitivity.

**ANALYSIS AND DISCUSSION**

**Quantitative Characterization of Forest Mosaics**

This analysis shows that eight of the 10 SIs selected can detect differences among different forest mosaics. Patch Richness Density, Patch Density, Edge Density, Landscape Shape Index, Area-Weighted Mean Shape Index, Patch Cohesion Index, Patch Area Density Mean, and Mean of Core Area are appropriate for characterizing different forest mosaics for at least one of the scales (window sizes) used. The determination of SIs able to effectively quantify the patch proximity and landscape fragmentation, or the optimal conditions to use them has not been attained. Nevertheless, our results show that the use of a combination of different SIs, used under optimal conditions (landscape scales), should allow the spatial and quantitative characterization of forest mosaic associated with a territory.

This study has revealed several points. 1) SIs must be used at the scale of the ecological phenomenon understudy. Empirical studies such as this one can serve as a reference.
Application: Use of SIs to Predict Key Wood Procurement Parameters

The analysis of the use of SIs to predict key wood procurement parameters showed that SIs could be used to predict procurement cost at a strategic level of planning. However, SIs do not provide an advantage over a more traditional method using GIS and road planning at the operational level.

---

Table 2—Results for the three procurement areas and the global model, from the predictive model produced for each parameter using SIs.

<table>
<thead>
<tr>
<th>Selected variables</th>
<th>Bowater</th>
<th>Smurfit-Stone</th>
<th>Abitibi-Consolidated</th>
<th>Global (3 territories)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Territories</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction + Access Road Maintenance</td>
<td>PD; AREA_MN</td>
<td>AREA_MN; PD</td>
<td>AREA_MN; PD</td>
<td>PD; AREA_AM</td>
</tr>
<tr>
<td></td>
<td>$r^2=0.52$</td>
<td>$r^2=0.66$</td>
<td>$r^2=0.29$</td>
<td>$r^2=0.66$</td>
</tr>
<tr>
<td>Construction + Extraction Road Maintenance</td>
<td>CA; PD</td>
<td>CA; PD</td>
<td>CA; PD</td>
<td>CA; PD</td>
</tr>
<tr>
<td></td>
<td>$r^2=0.92$</td>
<td>$r^2=0.63$</td>
<td>$r^2=0.39$</td>
<td>$r^2=0.741$</td>
</tr>
<tr>
<td>Transport Access Road</td>
<td>vol_dist_acc (ENN)</td>
<td>dist_vol_acc (ENN)</td>
<td>dist_vol_acc (ENN)</td>
<td>dist_vol_acc (ENN)</td>
</tr>
<tr>
<td></td>
<td>$r^2=0.99$</td>
<td>$r^2=0.55$</td>
<td>$r^2=0.70$</td>
<td>$r^2=0.82$</td>
</tr>
<tr>
<td>Transport Extraction Road</td>
<td>dist_vol_rec (ENN)</td>
<td>dist_vol_rec (ENN)</td>
<td>dist_vol_rec (ENN)</td>
<td>dist_vol_rec (ENN)</td>
</tr>
<tr>
<td></td>
<td>$r^2=0.95$</td>
<td>$r^2=0.94$</td>
<td>$r^2=0.94$</td>
<td>$r^2=0.91$</td>
</tr>
<tr>
<td>Forwarding distances</td>
<td>MIN; MAX; ED</td>
<td>MIN; MAX; ED</td>
<td>MIN; MAX; ED</td>
<td>MAX; MIN; ED</td>
</tr>
<tr>
<td></td>
<td>$r^2=0.94$</td>
<td>$r^2=0.93$</td>
<td>$r^2=0.95$</td>
<td>$r^2=0.94$</td>
</tr>
</tbody>
</table>

Table 3—For each factor impacting wood procurement cost, function developed with the use of SIs and the $R^2$.

<table>
<thead>
<tr>
<th>Factor impacting wood procurement cost</th>
<th>Function developed with the use of SIs</th>
<th>$R^2$(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access roads (figure 2-A)</td>
<td>Total Distance (TD)</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>$TD = 1.300 \times 10^8 \text{PD}^4 + 3.06 \times 10^5$</td>
<td></td>
</tr>
<tr>
<td>Extraction roads (figure 2-B)</td>
<td>Extraction Road Distance (ERD)</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>$ERD = 2.294 \times 10 \text{CA}^2 + 2.434 \times 10^7 \text{PD}^3 + 3.377 \times 10^2$</td>
<td></td>
</tr>
<tr>
<td>Forwarding distances (figure 2-C)</td>
<td>Average Distance (AD)</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>$AD = AD = 0.410 \text{MAX}^4 - 0.660 \text{MIN}^5 - 2.068 \times 10^4 \text{ED}^6 + 4.875$</td>
<td></td>
</tr>
<tr>
<td>Wood transport (access) (figure 2-D)</td>
<td>Wood Transport Access (WTA)</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>$WTA = 1.163 \times \text{ENN}^7 - 6.001 \times 10^3$</td>
<td></td>
</tr>
<tr>
<td>Wood transport (extraction) (figure 2-E)</td>
<td>Wood Transport Extraction (WTE)</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>$WTE = 1.267 \times \text{ENN}^8 - 3.891 \times 10^2$</td>
<td></td>
</tr>
</tbody>
</table>

$^1$ Patch Density of the harvesting area
$^2$ Core Area of the cut blocks
$^3$ Patch Density of the cut blocks
$^4$ Maximum distance between a line crossing the cut block major axis and the cut block edge
$^5$ Minimum distance between a line crossing the cut block major axis and the cut block edge
$^6$ Edge Density of the cut blocks
$^7$ Euclidean Nearest Neighbor between the harvesting sectors’ centroide and the mill
$^8$ Euclidean Nearest Neighbor between the harvesting sectors’ centroide and the access road

for determining it. 2) The scale must be constant when using SIs to compare landscapes. This implies that the landscapes must have the same cartographic scale, as well as the same resolution and extent. Moreover the landscapes should be generated according to the same standards. 3) SIs are not effective at all scales to detect different landscapes. In fact, they must be used under conditions (scales) that give them the most sensitivity. And 4), if in doubt, it is preferable to test SIs empirically in order to know their behavior.
Hence, at the strategic planning level, access roads, extraction roads, skidding distances, access for wood transport, transport for extraction of wood from harvesting sectors, and machinery displacement between blocks have been identified as the main spatial factors impacting the wood procurement cost (Figure 2). Of these factors, the regression analysis revealed that all but the displacement of machinery could use SIs to evaluate procurement cost. For the displacement of machinery, the SIs selected were not reliable to provide a robust evaluation.

A predictive model was produced for each parameter in each of the three procurement areas (Table 2). In all cases, the same SIs were identified as the most statistically significant. The predictive value is optimum when region specific models are used to predict the dependent variable. However, for all key parameters except “access road”, the global model can provide useful estimates to decision makers.

The function developed (using SIs) to quantify the key wood procurement parameters are presented in Table 3. This table also shows the R² value.

Hence, to determine wood procurement cost of the different factors impacting it, one can apply widely used cost functions and provide the estimated spatial parameter (Table 3).

**CONCLUSION**

In Quebec (Canada), the appearance of the “mosaic” forest management method, which aims at conciliating needs of different forest users by maintaining forest land integrity, imposes new spatio-temporal standards on forest resources managers. This trend of adapting forest operations to imitate the spatial pattern generated by the regime of natural disturbances and natural environment is increasingly popular in Quebec and elsewhere.

SIs are mathematical expressions having the goal of objectively quantifying aspects of landscapes’ spatial composition and configuration on categorical maps. Consequently, a combination of different SIs should enable the definition of forest mosaics’ spatial characteristics. It is from this premise that this study originated.

In this context, SIs can thus be used as a tool for forest planning in order to determine which management scenario will minimize the forest operation effects on a landscape’s structure. Moreover, they could be used, at a strategic resource management level, to evaluate wood procurement cost of a scenario or management method.

This study demonstrated that SIs could be part of a simple and robust methodology to spatially characterize forest mosaics and call attention to their possibilities for uses in forest resources management. It also showed the utilization potential of SIs to predict key wood procurement parameters and help evaluate wood procurement costs. Although preliminary results demonstrate that procurement cost predicted using an SI-based equation would only apply at a strategic planning level, and that results would be less accurate than those obtained with a more exact approach, significant time savings are possible nonetheless. Moreover, an SI-based approach may provide new methodological possibilities for advanced harvesting modeling.

**LITERATURE CITED:**


The Forest Service of the U.S. Department of Agriculture is dedicated to the principle of multiple use management of the Nation’s forest resources for sustained yields of wood, water, forage, wildlife, and recreation. Through forestry research, cooperation with the States and private forest owners, and management of the national forest and national grassland, it strives—as directed by Congress—to provide increasingly greater service to a growing Nation.

The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual’s income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA’s TARGET Center at (202) 720-2600 (voice and TDD). To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, D.C. 20250-9410, or call (800) 795-795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.