Geographic information systems (GIS) permit spatial analysis to be used to model forest management processes through a spatial decision support system (SDSS). However, these systems are often applied with little regard for the uncertainties in the supporting data, and that may have serious impacts on the SDSS outputs and the decisions made. Our understanding of the uncertainties inherent in spatial data has not kept pace with the analytical sophistication of the systems based on them.

Every data value, whether a single value or the mean of a set of values, is an estimate of the true value. The estimated variance of the values provides a measure of its uncertainty. The goal of uncertainty analysis is to estimate the distribution of results of some process, given input values and some measure of their uncertainty. With knowledge of the probability of various possible results of the decision-making process, more informed decisions can be made. Knowing the cost of uncertainty, forest managers can decide whether to improve data quality.

This paper presents a method of estimating the cost of using spatial data with known uncertainty through its application to a specific case of forest management decision-making. Specifically, known uncertainty was introduced into evaluation data used as part of a stumpage value model. Monte Carlo simulation was used to calculate the uncertainty in the resulting stumpage estimates. Decisions about which compartments to harvest were based on a simple ranking of the compartments based on their maximum potential stumpage (MPS) or profitability. MPS was calculated as the difference between the value of the products (sawlogs and pulpwood) at the mill and the costs to skid and haul the products to the mill. Skidding costs were calculated using a model that included slope, soil tractability, hydrology, roads, and compartment inventory as the input variables.

The study is based on the 3,000-acre, 97-compartment Heberg Memorial Forest, on the northern escarpment of the Allegheny Plateau in central New York. The terrain is moderately rugged, with elevations ranging more than 820 feet, and features a number of small streams and ponds and a road network. Land cover includes abandoned farmland, plantations of softwoods of various species, and second-growth northern hardwoods.

Methods

A timber-harvesting decision model. Harvest planning for forest products requires knowledge of the biological aspects and financial value of compartments and the geographic relationships between compartments. An SDSS is therefore an appropriate tool for making harvesting decisions (Reisinger and Davis 1987; Ferlow 1984; Osborne 1988, etc). Traditional timber supply analysis generalizes resource accessibility into a few broad zones based on distance from roads. The timber supply is calculated from values describing the typical timber composition within each accessibility zone. A limitation to this approach is that all trees are assumed to be transported straight to the nearest road regardless of any barriers to skidding (streams, lakes, cliffs) or terrain features that increase skidding costs (steep slopes, soils with poor tractability).

These features and other variables needed to compute costs can be represented in a raster GIS. This type of system represents thematic data as a matrix of values in cells similar in structure to that used to represent satellite images. The GIS has the ability to perform analyses both within a layer or image and between images. The GIS used is IDRISI (Eastman 1992), and the data layers in this system are called images.

IDRISI deals with barriers and areas of high skidding cost by computing an effective distance, that is, the distance to be traveled over a route that avoids barriers and areas of high cost if it is less costly to do so, for each cell in the image. Effective distance is calculated from a friction or resistance-to-movement image and can be expressed in units of distance, dollars, or gallons of fuel. Ferlow (1984) created the spatial model used to compute MPS. Liu (1994) implemented Ferlow’s model in IDRISI (Liu and Herrington 1993). The spatial model, shown in figure 1 (p. 28), can generally be expressed as

\[
MPS_s = \frac{\sum (MV_s - CL_s - CM_s)}{N_s}
\]

where summation is over all cells in a compartment (S = 1 to N) and \(MPS_s\) = average maximum potential stumpage ($ per acre) by compartment; \(MV_s\) = market value, which equals to volume per acre x mill price per species ($ per acre); \(CL_s\) = cost to landing, or skidding cost ($ per acre); \(CM_s\) = cost to mill, or hauling cost ($ per acre); and \(N_s\) = number of cells in the compartment.

Assumptions behind the MPS values so calculated include the following:

- Social factors play no role in harvest costs.
- All merchantable trees within a compartment are harvested.

Journal of Forestry 27
• Skidding is always to the nearest road over the least costly route.
• Sawlogs and pulpwood are hauled to mills in Cortland and Syracuse, New York, respectively.
• Fixed costs are not included.

DEM uncertainty. In this study only the elevation data or digital elevation model (DEM) had uncertainty; all other variables were considered free of uncertainty. Slope was calculated from the DEM, so it, too, has uncertainty. The accuracy of DEMs published by the US Geological Survey (1987) is reported as a root-mean-square-error (RMSE). The sources of uncertainty in DEMs are multitudinous (see Burrough 1986) but are assumed not to be due to mistakes or blunders; thus we use the term uncertainty here. USGS assumes that the error at any point is independent of that at any other point (that is, the error is not spatially autocorrelated). Some reports of independent or random errors in DEMs exist (Cáruso 1987; Carter 1989). Here we assume that errors in DEMs are normally distributed and that the RMSE is equivalent to the standard deviation of a sample from a normal distribution and is a measure of the uncertainty in the geographic data.

When uncertainty is added to a DEM by introducing random values to be added or subtracted from the true DEM, the resulting surface no longer looks like normal terrain. In fact, its spatial autocorrelation is reduced from ~0.95 to ~0.85. To correct for this, the perturbed DEMs used here were filtered with a $3 \times 3$ low-pass spatial filter. This effectively brought the spatial autocorrelation back to an acceptable level and the resulting surfaces again looked like natural terrain.

Monte Carlo simulation. In Monte Carlo simulation the input to a process is repeatedly perturbed so that a distribution of output results is obtained. In this case the DEM, and thus the derived slope image, was iteratively subjected to random perturbation to compute the distribution of MPS.

A computer program was used to generate multiple realizations of the process. Three levels of uncertainty, called cases, represented by random surfaces with a specified standard deviation $\sigma$, were used (table 1). The process was repeated 200 times to produce 200 independent scenarios with MPS values for each compartment. The number of iterations was chosen based on an analysis of the convergence of the mean MPS for each case with the true MPS as the number of iterations in the case was increased. These 200 sets of output values constitute a random sample of output derived from the probability distribution of the input. Figure 2 illustrates the process used to generate perturbed DEMs.

The IDRISI macro program for this process is outlined in the steps below (a macro is a DOS batch file using the IDRISI module’s command line options). Numbers in parentheses refer to images from previous steps, and brackets enclose IDRISI module names.

1. Generate an image of random values drawn from a normal distribution with mean 0 and standard deviation $\sigma$ [RANDOM].
2. Add (1) to the original DEM to produce a perturbed elevation image [OVERLAY].
3. Smooth (2) with a low-pass filter [FILTER].
4. Create a % slope image from (3) [SURFACE].
5. Add the friction values for slope (3), soil, and streams to produce a total friction image [OVERLAY].

6. Calculate the effective distance between each pixel and the road network using (5) and the road image [COST].

7. Generate a distance cost image in $ from (6), sawlog and pulpwood volume, and harvesting equipment costs [SCALE].

8. Calculate MPS for all pixels (equation 1) [OVERLAY].

9. Calculate the average MPS for each forest [EXTRACT].

10. Rank all compartments by their MPS value.

The cost of uncertainty for the harvesting decision process. The assumption made was that the introduced uncertainty in the DEM would affect the selection of the 10 most profitable compartments by changing their ranking. This was based on the fact that the slope values and hence also the cost derived from the DEM would increase with $ because there would be an increase in the number of cells with steeper slopes (fig. 3).

The expected cost of uncertainty (E\( [CU] \)) analysis (Morgan and Henrion 1990) determines how much money can be lost by basing the harvesting decision on data with uncertainty. E\( [CU] \) should increase as the uncertainty level in the data increases. The difference between E\( [CU] \) values for different levels of uncertainty provides an upper limit on what could be spent to reduce uncertainty from the higher to the lower levels. The expected cost of uncertainty can be expressed as

\[
E\left([CU]\right) = MPS_u - MPS_t \quad (2)
\]

where E[ ] means an expected value, \( [CU] \) is the expected cost of uncertainty, and the subscripts \( u \) and \( t \) refer to values calculated with the uncertain and true data, respectively. Substituting (1) into (2) gives

\[
E\left([CU]\right) = E\left[MV - CL_u - CM\right] - E\left[MV - CL_t - CM\right] \quad (3)
\]

To examine the differences between the harvesting decisions made at different levels of uncertainty, (3) can be expressed as follows:

\[
E\left([CU]\right)_c = MPS_t - \left(\sum E\left[MV - CL_{uc} - CM\right]\right)_c / n \quad (4)
\]

where \( n \) is the number of iterations in case \( c \). Since \( MV \) and \( CM \) are constants in this study, (4) becomes

\[
E\left([CU]\right) = CL_t - \left(\sum E\left[CL_{uc}\right]\right) / n \quad (5)
\]

Results and Discussion

We look first at the expected costs that might be realized from the use of uncertain data, and then at the effect of uncertainty on the decisions made.

Expected cost of uncertainty. The costs of uncertainty may arise from one of two sources or a combination: (1) lower appraised value of stumpage, with no change in compartment ranking; and (2) change in compartment ranking with no change in average compartment values over all compartments. The first is uninteresting for the
Table 2. Partial table of average cost of uncertainty (CU) and maximum potential stumpage (MPS) for the forest (in dollars).

<table>
<thead>
<tr>
<th>Monte Carlo iteration</th>
<th>True Case</th>
<th>Case 1, $\sigma = 4.0$ meters</th>
<th>Case 2, $\sigma = 5.4$ meters</th>
<th>Case 3, $\sigma = 7.0$ meters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MPS</td>
<td>MPS</td>
<td>MPS</td>
<td>MPS</td>
</tr>
<tr>
<td>Mean</td>
<td>$18,730.83$</td>
<td>$1,559.00$</td>
<td>$1,980.00$</td>
<td>$2,458.00$</td>
</tr>
<tr>
<td>1</td>
<td>$18,730.83$</td>
<td>$17,179.98$</td>
<td>$1,556.97$</td>
<td>$16,742.45$</td>
</tr>
<tr>
<td>2</td>
<td>$18,730.83$</td>
<td>$17,176.28$</td>
<td>$1,554.55$</td>
<td>$16,741.98$</td>
</tr>
<tr>
<td>3</td>
<td>$18,730.83$</td>
<td>$17,179.04$</td>
<td>$1,551.79$</td>
<td>$16,752.64$</td>
</tr>
<tr>
<td>4</td>
<td>$18,730.83$</td>
<td>$17,172.49$</td>
<td>$1,558.34$</td>
<td>$16,778.49$</td>
</tr>
<tr>
<td>5</td>
<td>$18,730.83$</td>
<td>$17,163.74$</td>
<td>$1,567.09$</td>
<td>$16,769.10$</td>
</tr>
<tr>
<td>6</td>
<td>$18,730.83$</td>
<td>$17,178.24$</td>
<td>$1,562.59$</td>
<td>$16,740.38$</td>
</tr>
<tr>
<td>7</td>
<td>$18,730.83$</td>
<td>$17,180.08$</td>
<td>$1,550.75$</td>
<td>$16,761.67$</td>
</tr>
<tr>
<td>8</td>
<td>$18,730.83$</td>
<td>$17,191.20$</td>
<td>$1,539.63$</td>
<td>$16,741.46$</td>
</tr>
<tr>
<td>9</td>
<td>$18,730.83$</td>
<td>$17,166.38$</td>
<td>$1,564.45$</td>
<td>$16,756.29$</td>
</tr>
<tr>
<td>10</td>
<td>$18,730.83$</td>
<td>$17,168.00$</td>
<td>$1,562.83$</td>
<td>$16,760.88$</td>
</tr>
<tr>
<td>11</td>
<td>$18,730.83$</td>
<td>$17,186.09$</td>
<td>$1,550.75$</td>
<td>$16,778.49$</td>
</tr>
<tr>
<td>12</td>
<td>$18,730.83$</td>
<td>$17,195.16$</td>
<td>$1,539.63$</td>
<td>$16,741.46$</td>
</tr>
<tr>
<td>13</td>
<td>$18,730.83$</td>
<td>$17,165.58$</td>
<td>$1,562.59$</td>
<td>$16,761.67$</td>
</tr>
<tr>
<td>14</td>
<td>$18,730.83$</td>
<td>$17,165.19$</td>
<td>$1,551.79$</td>
<td>$16,756.29$</td>
</tr>
<tr>
<td>15</td>
<td>$18,730.83$</td>
<td>$17,165.58$</td>
<td>$1,562.59$</td>
<td>$16,761.67$</td>
</tr>
</tbody>
</table>

Note: The table shows the results for only 15 iterations. Means are for all 200 iterations.

For purposes of improving GIS data, but the second would imply that prioritization to arrange harvesting in order of decreasing stumpage values would benefit from better elevation data.

For each of the 200 iterations in the Monte Carlo process in each case, a single $MPS_u$ and $CL_u$ values for each compartment was produced. The 10 compartments with the highest $MPS_u$ were ranked in descending order. These compartments represented the optimum harvesting decision for each case. Equation (5) was then used to calculate the $E[C_U]$ for the best harvesting decision. Table 2 shows examples of $MPS_u$ and $CL_u$ for the three simulated uncertainty levels (cases 1, 2, and 3). The average $E[C_U]$ values for each case (equation 5) are shown in table 2, and probability distribution for $C_U$ for case 3 is plotted in figure 4 (p. 29).

The results indicated the following:
1. The total $E[C_U]$ value for case 1 was $1,558$, which is about 8.3% of expected total $MPS$; for case 2, $1,980$ and 10.65%; and for case 3, the highest level of uncertainty, $2,458$ and 13.1%.
2. The $MPS_u$ values derived from images with spatial uncertainty were less than those for the true image, and this difference increases with increasing uncertainty.
3. The probability distributions for the three levels of uncertainty are nearly normal.

Effect on decision making. It is clear that there is a cost associated with the use of uncertain data. As more uncertainty was introduced into the DEM, maximum and mean values of slope derived from these data increased. Since skidding cost increases with slope, the MPS values decrease with increasing uncertainty. Thus, the more uncertainty in the spatial data, the more serious the impact on decisions.

Were there any differences in the decisions made about harvesting using the simple ranking decision process? Figure 5 shows the frequency distribution of the rankings of the top 13 most profitable compartments for case 3 and lists the ranking of the compartments for the true case.

Clearly, there were some differences, since the list for case 3 includes three additional compartments. It is also clear that the same ranking was not made all the time. However, the probability that the uncertain and true rankings would be...
the same is high. The probability of making a different decision is finite—but is it significant? And what is penalty for making the wrong decision? That is for the forest manager to decide.

Conclusions

A method for examining the effects of uncertainty in spatial databases on forest harvest decisionmaking was described. The expected cost of uncertainty was used to estimate the cost of using uncertain elevation data for forest harvest decisions. The results from the analysis indicate that uncertain elevation data used as input to a harvesting SDSS could yield uncertainties in cost of up to 13% of true MPS value. As a result of uncertain data, the probability distribution of the ranking order of the compartments could vary. However, the simple ranking model used for this process is not very sensitive to even the highest level of uncertainty in the data.

Even though the model and the decision process were not highly sensitive to the simulated error in elevation, the results suggest that we pay attention to data quality. The ability to examine how various sources of uncertainty contribute to the model outputs is potentially of great value for resource managers who need reliable information for realistic decisions. In real situations other models and decision processes may be much more sensitive to the errors in elevation or other data than the model used here. For a specific SDSS, the methods reported here can be used to assess the costs of using data with different levels of uncertainty and the sensitivity of the decision process to that uncertainty. It does appear that in forest resources management, attention to data quality can pay off.

Literature Cited


Eastman, J.R. 1992. IDRISI, Version 4.0, Clark University, Graduate School of Geography, Worcester, MA.


About the Authors

Rei Liu is GIS coordinator and senior environmental engineer, Virginia Department of Forestry, PO Box 3758, Charlottesville 22903; Lee P. Herrington is professor and director of resources information management and applied GIS at State University of New York, College of Environmental Science and Forestry, Syracuse. The authors are grateful to J. Toll for his comments.

MAPS FOR YOUR TOTAL RESOURCE PLAN

British Columbia's new FOREST PRACTICES CODE requires better mapping—for VQOs, biodiversity, watershed restoration and inventory management.

MAPS 3D software, working with MicroStation PC®, will provide you with mapping capability to match the new demands.

Already in use throughout BC, MAPS 3D can help you too.

Contact: Forestry Spatial Data Group
Pacific International Mapping Corp.
101-4218 Commerce Circle, Victoria BC Canada V8Z 6N6
Tel: (604) 727-0727 Fax: (604) 727-3153