For. Sci. 60(1):137–148
http://dx.doi.org/10.5849/forsci.12-090

operations research

A Robust Model for Protecting Road-Building and Harvest-Scheduling Decisions from Timber Estimate Errors

Cristian D. Palma and John D. Nelson

Road-building and harvest-scheduling decisions are primarily based on timber estimates and forecasts that are known to contain errors. It has been shown that in the presence of constraints, decisions generated under these conditions are likely to become infeasible. Therefore, solutions are required that can ensure constraint fulfillment despite the estimation errors. We present a robust model formulation of a multiperiod road-building and harvest-scheduling problem in which protection against minimum demand infeasibility is sought despite the existence of timber estimates that are defined as continuous ranges of values instead of point estimates (as is usually the case in this type of problems). We compare the benefits of this robust formulation with those of the traditional deterministic option and explore the tradeoff between the robustness of the solutions and its impact on the objective function. By simulating different scenarios of the timber coefficient realizations, it is shown that the robust approach produces solutions that are less sensitive to errors in the timber estimates at the expense of a slight reduction in the objective function.

Keywords: road building, uncertainty, robust optimization, timber estimate errors, forest planning

Medium-term tactical decisions that cover a 2- to 5-year planning horizon typically involve harvest-scheduling decisions and road construction. The difficulty of solving this problem has led researchers to propose different mixed binary models and to apply various solution approaches (Weintraub and Navon 1976, Guignard et al. 1998, Richards and Gunn 2000, Andalaft et al. 2003) to find (near) optimal solutions to these models. Most likely because of this complexity, the models assume that the data are perfectly known and ignore the inherent uncertainty in the coefficients of the model. Although uncertainty has been traditionally associated with a lack of numerical information to describe the future, unlike risk that refers to a quantified uncertainty (Davis and Johnson 1987), we use uncertainty in this work simply to indicate lack of certainty, whether measurable or not.

This uncertainty may originate from a variety of sources that are well described in the literature (Marshall 1987, Mowrer 2000, Regan et al. 2002). Biological processes that are not completely understood, natural disasters, and changes in social and economic conditions, in combination with long time horizons, all affect the future consequences of forest-planning decisions made in the present. In addition, uncertainties may result from the process of estimating current and future resource levels. Different sampling methods used to determine the stand volume estimates will report different estimate errors of the current inventory, and statistical models used to project this inventory will also contribute to this inexactness. Although an extensive area may produce a diversification effect that could reduce the effects of some of these uncertainties, the decisionmaking process still must be carried out in an uncertain environment.

In the context of the sequential decisionmaking that takes place in real-world systems, there is potential recourse if the forecasted outcomes are not realized, and the consequences may not necessarily be large. Because many of the real system decisions are guided by model forecasts, it is possible to gain an understanding of the uncertainty in the real system by examining the feasibility of the models for the cases in which information does change. Solutions from deterministic models are likely to become infeasible if evaluated with observed data. Although decisions implemented in the present might appear optimal, these decisions will most likely become suboptimal and mathematically infeasible once the uncertainties are realized. The information used to make these decisions will not necessarily be observed in actuality (Hof et al. 1988, Pickens and Dress 1988). This situation renders the search for good quality and stable solutions (rather than strictly optimal solutions) highly

Manuscript received August 1, 2012; accepted March 13, 2013; published online July 4, 2013.

Affiliations: *Cristian D. Palma (cristianpalma@ingenieros.udd.cl), Universidad del Desarrollo, Faculty of Engineering, Concepcion, Chile. John D. Nelson (john.nelson@ubc.ca), Department of Forest Resources Management, University of British Columbia., Vancouver, Canada.*

Acknowledgments: *This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada. We thank the editor, an associate editor, and two anonymous reviewers for their comments and suggestions that helped us to improve the original version of this article.*

relevant when decisions are implemented. In this article, we address a harvest-scheduling and road-building problem in which the timber volume estimates are uncertain and are defined by a range of values rather than by a point estimate. We use robust optimization to identify solutions that remain highly feasible and close to the optimal solution even if any combination of the possible timber values is observed in actuality.

Approaches for Dealing With Uncertainty

Stochastic programming (SP), chance-constraint programming (CCP), and fuzzy set theory (FS) are the best-known techniques that include uncertainty in the optimization models. By considering a set (usually discrete) of scenarios of random events or uncertainties, SP allows the identification of a good solution for all scenarios or for the most likely scenarios. The greatest disadvantage of this technique is that as the number of scenarios increases, the models become computationally difficult to solve and solution methods such as decomposition and statistical approximation must be applied (Birge and Louveaux 1997). In addition, the probability distribution of the scenarios is required but is generally not known.

In CCP, constraints with at least one random coefficient are modeled as probabilistic statements. The probability distributions of uncertain coefficients are assumed to be known and the constraints are required to be met with a minimum probability that is exogenously determined. Although particular cases of CCP models are easy to solve (e.g., when only the right-hand side of a constraint is uncertain), the models become nonlinear in most cases (Kall and Wallace 1994, Birge and Louveaux 1997). Finding exact solutions is difficult (Chen et al. 2007), thus motivating the search for approximate solution techniques (Birge and Louveaux 1997).

FS provides a different approach to dealing with certain types of uncertainty by defining a degree of membership for a parameter to a set of possible values. The more likely the value of a parameter, the greater is its degree of membership. This concept is used to model the objectives and constraints for which a combination of their degree of membership can be maximized (Zimmermann 1996). This approach may be appropriate if there is vagueness, for example, in the meaning of certain events, phenomena, or statements such as the preferences among different objectives or the definition of aspirational goals. However, the appropriateness of the approach is not clear if the model represents a lack of information with respect to the value of the parameters. In the latter case, FS actually results in a relaxed version of the traditional deterministic problem in which constraint violations are allowed and better objective functions are obtained as a consequence.

These three techniques have been used in harvest-scheduling problems. For example, SP has been used to address the uncertainty in timber yield (Hoganson and Rose 1987, Eriksson 2006) and fire losses (Gassmann 1989, Boychuk and Martell 1996), and uncertainty in timber yield has also been considered using CPP (Weintraub and Vera 1991, Pickens et al. 1991, Hof et al. 1992, Weintraub and Abramovich 1995). Most recently, CCP with approximation solution techniques were applied to a fire budget allocation problem with uncertain fire suppression costs and to a habitat restoration problem in which the survival parameters were uncertain (Bevers 2007). With use of FS, the uncertainties in timber yield (Bare and Mendoza 1992), periodic harvesting goals (Hof et al. 1986, Pickens and Hof 1991), and vagueness in seral-class definitions (Boyland et al. 2006) have been considered in harvest-scheduling problems. In the context of multiobjective models, fuzziness in the objective function coefficients (Mendoza et al. 1993, Stirn 2006), as well as in the goal definitions (Ells et al. 1997, Maness and Farrell 2004), has been addressed using FS.

Uncertainty in road-building decisions has been scarcely addressed, most likely because of the additional difficulty in solving integer models. To our knowledge, only a few applications that address the uncertainty in road-building problems have been reported, and all of them use SP. In Olsson (2007), road upgrade decisions with an uncertain length for the period of poor road conditions were examined, and in Alonso-Ayuso et al. (2009), the harvest and road construction problem was considered for cases in which different price scenarios were defined. In both cases, the explicit consideration of the uncertainties was recommended over the traditional deterministic model because the decisions performed better in terms of their stability. The reduced number of scenarios that the methodology can handle was also clear in these works.

The aims of this article are two-fold. First, we explore the effect of random volume uncertainties on harvest-scheduling and road-building decisions, and second, we do this using robust optimization (RO), a mathematical programming approach that has not yet been extensively applied in forest resources management and that allows finding decisions that are less sensitive to uncertain parameters. This method has been applied in engineering (Ben-Tal and Nemirovski 2002), network design (Bertsimas and Sim 2003, Ordonez and Zhao 2007), and inventory theory (Bertsimas and Thiele 2006), among other problems. In natural resources planning, applications of RO can be found in water supply management (Chung et al. 2009) and strategic forest harvest-scheduling problems (Palma and Nelson 2009, 2010).

The RO Approach

We base our model development in the RO approach proposed by Bertsimas and Sim (2004). Other approaches have also been proposed (Mulvey et al. 1995, El Ghaoui et al. 1998, Ben-Tal and Nemirovski 2000) but have resulted in models that are computationally more complex. The approach used in this work produces models of the same type as the original model. That is, if the original model is a mixed-integer linear model, then the robust model is also mixed-integer linear, therefore eliminating the need for specific solution techniques other than that used to solve the original problem.

The RO approach operates as a buffering strategy. An additional term, known as the protection function, is added to each constraint for which feasibility is highly desirable and for which the technical coefficients are uncertain. Let us consider the general lower bound constraint, in which $i = 1, \ldots, m$ and $j = 1, \ldots, n$

$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i \quad \forall i \tag{1}$$

where a_{ij} are assumed as uniformly distributed in the range $[\bar{a}_{ij} - \hat{a}_{ij}, \bar{a}_{ij} + \hat{a}_{ij}]$ with $\bar{a}_{ij}$ as the coefficient estimate and $\hat{a}_{ij}$ as the accuracy of the estimate. Although other distributions might be more realistic, the assumption of a uniform distribution of errors allows robust models to remain linear. If, for each a_{ij}, a scaled deviation is defined as $\nu_{ij} = (a_{ij} - \bar{a}_{ij})/\hat{a}_{ij}$, then $\nu_{ij} \in [-1, 1]$, and for each constraint, $\sum_{j=1}^{n} \nu_{ij}$ can theoretically take on values between $-n$ and n. We can now limit the number of uncertain coefficients that are allowed to change in constraint i by considering $\sum_{j=1}^{n} |\nu_{ij}| \leq \Gamma_i$ where Γ_i is known as the protection level of constraint i. Three possibilities can be identified: if $\Gamma_i = 0$, the ν_{ij} for all j in constraint i are forced to 0 such that $a_{ij} = \bar{a}_{ij}$ for all j, and constraint i has no protection against uncertainty; if $\Gamma_i = n$, constraint i is totally protected against uncertainty because all ν_{ij} in constraint i are allowed to take on a nonzero value; and if $0 < \Gamma_i < n$, constraint i is partially protected against uncertainty, in which case a subset of the ν_{ij} can be different from 0. The robust version of Equation 1 is then

$$\sum_{j=1}^{n} \bar{a}_{ij} x_j - \beta_i(x^*, \Gamma_i) \geq b_i \quad \forall i \tag{2}$$

The protection function of constraint i, $\beta_i(x^*, \Gamma_i)$, depends on the user-defined protection level Γ_i and is specific for a given solution vector x^*. RO looks for the optimal buffer that maintains the left-hand side of Equation 2 $\geq b_i$ for different values of a_{ij}. For a given solution vector, the protection function determines the buffer required if up to Γ_i parameters are allowed to change; i.e., the buffer size corresponds to the sum of the Γ_i largest deviations produced by a fixed value of x. In other words, the protection function equals the objective function of the following optimization problem

$$Max \sum_{j=1}^{n} \hat{a}_{ij} x_j^* w_{ij} \tag{3}$$

subject to

$$\sum_{j=1}^{n} w_{ij} \leq \Gamma_i \tag{4}$$

$$0 \leq w_{ij} \leq 1 \; \forall j \tag{5}$$

where w_{ij} are new variables that represent the random variable of the scaled deviation ν_{ij} described above. Although this model is linear for the given solution vector x^*, this is not the case when x is variable. However, its dual can be used to express this function linearly in 2. If z_i and p_{ij} are the dual variables of the constraints 4 and 5, respectively, then the dual of 3–5 is

$$Min \; \Gamma_i z_i + \sum_{j=1}^{n} p_{ij} \tag{6}$$

subject to

$$z_i + p_{ij} \geq \hat{a}_{ij} x_j^* \; \forall j \tag{7}$$

$$p_{ij} \geq 0 \; \forall j \tag{8}$$

$$z_i \geq 0 \tag{9}$$

Because model 3–5 is feasible and bounded for all $\Gamma_i \in [0, n]$, then its dual, model 6–9, is also feasible and bounded, and their objective functions coincide (Bertsimas and Sim 2004). Therefore, $\beta_i(x^*, \Gamma_i)$ is equal to the objective function of model 6–9 and can be substituted into Equation 2 to produce its robust version.

$$\sum_{j=1}^{n} \bar{a}_{ij} x_j - \Gamma_i z_i - \sum_{j=1}^{n} p_{ij} \geq b_i \; \forall i \tag{10}$$

$$z_i + p_{ij} \geq \hat{a}_{ij} x_j \quad \forall i, j \tag{11}$$

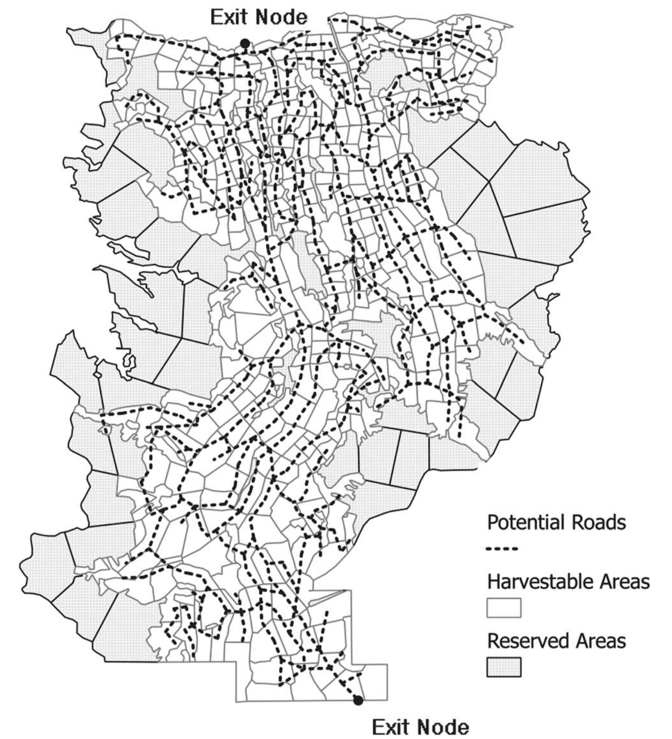

Figure 1. A harvestable area of 7,554 ha was used as the study area and included 431 stands and 412 potential roads.

$$p_{ij} \geq 0 \quad \forall i, j \tag{12}$$

$$z_i \geq 0 \quad \forall i \tag{13}$$

Methods

In this section, we describe the study area, both the deterministic and robust formulations, the data used, and the simulation experiments performed to test the resulting infeasibility rates.

Study Area

In this work, we determined the road and harvest decisions in a 11,675-ha area located on mid-Vancouver Island, British Columbia, Canada, of which 7,554 ha were available for harvesting (Figure 1). The area is divided into 431 harvestable stands with an average size of 17.5 ha and ranging from 1.6 to 52.3 ha. We considered 412 potential road segments and two demand nodes for which a minimum timber supply must be guaranteed over three planning periods. To cover the planning horizon usually considered in tactical decisions while simultaneously providing additional detail in the first periods, the periods were defined as 2, 3, and 4 years, respectively, which correspond to a 9-year planning horizon. Because a considerable portion of this area could be harvested in a relatively short period of time, 15% of the volume is to be retained when a stand is harvested. Inventory data and harvesting and transportation costs are available for the area.

Harvest-Scheduling and Road-Building Models

The road-building problem is a network design problem in which the arcs represent segments of potential roads and the nodes represent either timber sources or the intersection of the arcs. Roads must be built and harvesting decisions (as well as timber flow) must be determined to transport the harvested timber toward the exit nodes that connect the forest to the demand nodes. In our formulation, we minimized the discounted harvesting, transportation, and road-building costs subject to a minimum timber demand. Adjacency constraints were not included, and only one timber product and one road standard were assumed. Although external timber supply is a common practice used to meet timber demand constraints in this type of problem, this option was omitted to facilitate the analysis of the effect of infeasibilities of these constraints. In other words, if the external timber supply were possible, any need for extra production due to an uncertain future would be solved by buying the extra timber in the market and not changing the harvest decisions.

Because integrated harvest-scheduling and road-building models are difficult to solve to optimality, various solution approaches have been proposed to identify good solutions, e.g., mixed-integer solvers (Weintraub and Navon 1976), heuristics and meta-heuristics (Weintraub et al. 1995, Clark et al. 2000, Richards and Gunn 2000), and Lagrangean relaxation (Andalaft et al. 2003). In addition, strengthening and lifting techniques are also simple steps used to facilitate the model solution. By adding logical inequalities (or "triggers") and increasing the dimension of a constraint space (or "lifting"), the solution space of the LP relaxation is reduced, and therefore fewer solutions are evaluated and better bounds for the integer solution techniques can be obtained (Guignard et al. 1998). We opted for a strengthened formulation of the problem and solved it using commercial optimization software as detailed later.

The following nomenclature will be used in both the deterministic and robust formulations.

Sets:

Nodes in the network (supply, intersection and exit nodes) ($i = 1, \ldots, I$; $j = 1, \ldots, J$)
i, j

Demand node ($n = 1, \ldots, N$)
n

Potential road corresponding to an undirected arc connecting two nodes; i.e., each road r supports timber flow on directed arcs ij and ji ($r = 1, \ldots, R$)
r

Stand ($s = 1, \ldots, S$)
s

Time period ($t = 1, \ldots, T$).
t

Set of stands that supply timber to node i
$s(i)$

Coefficients and Parameters:

Minimum demand in node n in period t (m^3)
d_{nt}

Total volume in stand s in period t (m^3)
v_{st}

Cost of building road r (\$)
cr_r

Cost of transportation between nodes i and j (or demand node n) in period t (\$/m^3)
$ct_{ij(n)t}$

Total cost of harvesting stand s in period t (\$)
ch_{st}

Discount factor applied to period t
$\propto_t$

Decision Variables:

Binary variable that represents whether road r is built (1) in t or not (0).
X_{rt}

Binary variable that represents whether stand s is harvested (1) in t or not (0).
Y_{st}

Timber flow (directed) between nodes i and j in period t [m^3].
F_{ijt}

Timber flow (directed) from node i to demand node n in period t [m^3]
F_{int}

Deterministic Model

The sum of the road building, harvesting, and transportation costs was minimized as follows

$$Min \sum_{t=1}^{T} \propto_t \left(\sum_{r=1}^{R} cr_r X_{rt} + \sum_{s=1}^{S} ch_{st} Y_{st} + \sum_{i=1}^{I} \sum_{j=1}^{J} ct_{ij} F_{ijt} + \sum_{i=1}^{I} \sum_{n=1}^{N} ct_{in} F_{int} \right) \quad (14)$$

Subject to the following set of constraints

$$\sum_{t=1}^{T} Y_{st} \leq 1 \quad \forall s \quad (15)$$

$$\sum_{s \epsilon s(i)} v_{st} Y_{st} + \sum_{j=1}^{J} F_{jit} = \sum_{j=1}^{J} F_{ijt} + \sum_{n=1}^{N} F_{int} \ \forall i,\ t \quad (16)$$

$$\sum_{i=1}^{I} F_{int} \geq d_{nt} \ \forall n,\ t \quad (17)$$

$$\sum_{k=1}^{t} (F_{ijk} + F_{jik}) \leq M_{rt} \sum_{k=1}^{t} X_{rk} \ \forall (i, j) \ \epsilon \ r,\ t \quad (18)$$

$$Y_{st} \leq \sum_{k=1}^{t} \sum_{r \epsilon r(s)} X_{rk} \ \forall s,\ t \quad (19)$$

Constraint 15 limits the number of times that each stand can be harvested during the planning horizon. Constraint 16 imposes the conservation of flow at each node and within each period. The inflow to a node (left-hand side) may come directly from harvesting stands or from other nodes, whereas the outflow (right-hand side) may go to other nodes in the network or to the demand nodes.

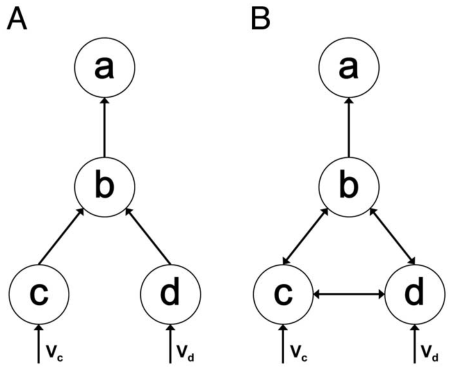

Figure 2. In a tree structure: (A) only one flow direction is possible so that the maximum flow in an arc can be determined as a cumulative flow (Max Flow$_{cb}$ = V_c; Max Flow$_{ba}$ = $V_c + V_d$). If cycles are present: (B) the maximum flow in an arc is more difficult to determine because complex flow interactions might result (Andalaft et al. 2003).

Constraint 17 defines the lower bound of the timber transported to the demand node m in each period. Constraints 18 and 19 correspond to the road construction trigger (lifted with respect to time) and a logical inequality, respectively; these both contribute to a strong formulation of the harvesting and road-building problem (Andalaft et al. 2003). In 18, the timber flow in any arc and up to a given period is only possible if the corresponding road has been built in the same or in a previous period. M_{rt} is a large number that allows the flow to be greater than 1. Because its value has a major impact on the solution process of the model, the lowest possible value that preserves the original integer optimum should be used (Andalaft et al. 2003). When maximum demands are present, the sum of all maximum demands represents an upper bound for the timber flow in each period and thus can be used as a value for M in the corresponding period. If there are no maximum demand constraints, as in our case, the total volume of the forest in each period, $\Sigma_s v_{st}$, is another option. However, because a minimum demand is required in each period, not all of the forest can be harvested in the same period because a certain amount of timber must be retained to meet demands in other periods. Therefore, a better upper bound for the flow for every arc and period t would be $\sum_s v_{st} - \sum_{l \neq t} d_l$. Moreover, the tightest upper bound can be determined if the arcs have a treelike structure (Figure 2). Unlike the cycle structures, only one flow direction is possible in treelike structures, and the potential flow in an arc cannot be larger than the maximum cumulative production of the source nodes from the leaves to the tree base.

For each arc and period, the flow upper bound was then determined as follows

$$M_{rt} = \begin{cases} M_{\underline{r}t} + \sum\limits_{s \epsilon s(i),(i,j) \epsilon r} v_{st} & \text{if r is in a tree structure} \\ \sum\limits_{s=1}^{s} v_{st} - \sum\limits_{l \neq t} d_l & \text{if r is in a cycle structure} \end{cases}$$

where $\underline{r}$ is the preceding arc of r from the leaves to the tree base and node i is the node in r closest to the leaf. $M_{\underline{r}t} = 0$ for terminal nodes. These calculations were automatically computed.

Constraint 19, known as a project-to-road trigger (Andalaft et al. 2003), allows harvesting of a stand in a period only if at least one road connecting the stand to the road network has been built in the same or in a previous period. In this equation, $r(s)$ represents the set of roads that connects stand s to the road network.

Robust Model

To protect the minimum demand constraint from infeasibilities, Equation 17 was modified, and new variables and constraints were added as described previously. We assumed that the real volume, v_{st}, belongs to and is symmetrically distributed in the range $[\bar{v}_{st} - \hat{v}_{st}, \bar{v}_{st} + \hat{v}_{st}]$, where $\bar{v}_{st}$ is the volume estimate and $\hat{v}_{st} = \bar{v}_{st} \cdot e_t$ is its estimate error.

A special feature of this problem deserves attention. The network structure of the problem makes it such that uncertain coefficients (v_{st}) are not included in the equation to be buffered, i.e., Equation 17. In fact, these coefficients do not depend on the demand at node n. Therefore, the protection function β_t will not provide a harvest buffer for each node n but only a total buffer at the period level. However, we can assign a portion of this total buffer to each node n by introducing the parameter δ_{nt}, with $\Sigma_n \delta_{nt} = 1$. In our case, both demands were assumed to be equally important in terms of the need for reducing infeasibilities, and thus δ_{nt} took on the value of the proportion of the demand that each destination represents out of the total demand (i.e., 0.6 for node 1 and 0.4 for node 2 for all periods). Although only flow variables (F_{int}) are present in the constraint to be buffered, these variables are linked to the harvest variables (Y_{st}), which are associated with the uncertain coefficients, through constraint 16. The additional timber flow needed to satisfy the demand constraint (Equation 17) forces the additional harvest (Equation 16) and the required road construction (Equation 18).

Constraint 17 of the original model was subsequently replaced by the following set of constraints

$$\sum_{i=1}^{I} F_{int} - \delta_{nt}\left(\Gamma_t z_t + \sum_{s=1}^{S} p_{st}\right) \geq d_{nt} \ \forall n, t \tag{20}$$

$$z_t + p_{st} \geq \hat{v}_{st} Y_{st} \ \forall s, t \tag{21}$$

$$p_{st} \geq 0 \ \forall s, t \tag{22}$$

$$z_t \geq 0 \ \forall t \tag{23}$$

where z_t and p_{st} are new variables. All other equations of the original model remain the same. The models were implemented in OPL Development Studio 5.2 (CPLEX 10.2.0 as optimizer) on an Intel Core 2 Quad 2.5 GHz computer with 4 GB of RAM.

Uncertain Coefficients and Model Parameters

A stable annual timber supply requirement was determined in such a way that approximately 70% of the total area needs be harvested during the planning horizon. Sixty percent of this supply requirement was arbitrarily assigned to one of the two demand nodes, and the rest was assigned to the other demand node. We note that other demand configurations do not affect the methodology and the conclusions of the study, although specific decisions may change. Demand levels for each period, therefore, resulted in 396,000, 594,000, and 792,000 m^3 for demand node 1, and 264,000, 396,000, and 528,000 m^3 for demand node 2. The annual discount rate was assumed to be 4%.

The timber forecast was based on both the current volume estimate and on a growth-and-yield model that projected this volume to future periods. The error of the initial volume estimate, e_1, was used to define two scenarios that might represent two sampling intensities. We assumed $e_1 = 10\%$ in the first scenario (referred to as SCE10) and $e_1 = 15\%$ in the second scenario (referred to as SCE15). In both cases, we assumed that errors increased 1% annually, which reflects the increasing uncertainty observed as the estimate goes farther into the future. Assuming that calculations occur at the midpoint of each period, the errors for the second and third periods were 12 and 16% in SCE10 and 17.5 and 21% in SCE15, respectively. The volume of a stand s in period t was therefore assumed to be random and uniformly distributed in the range $\bar{v}_{st} \pm \bar{v}_{st}e_t$ as a result of unbiased errors in the initial estimation of the stand area or current volume or in the forecast model. Because uncertainties have to be noncorrelated in the same constraint (i.e., the volume must be independent within the same period), spatial independence was also assumed.

As mentioned previously, the degree of conservatism in satisfying a constraint is controlled by the user-defined protection level, Γ_t. The immediate question that arises is how large should this parameter be to get a desired feasibility rate. Although there is no exact expression that will output this parameter, there exist bounds that relate a desired probability of the constraint violation to the protection level required (e.g., Bertsimas et al. (2004), Bertsimas and Sim (2004)). However, as noted in Palma and Nelson (2009), these bounds represent only a weak estimate of this probability and therefore overestimate the protection level and excessively affect the objective function value. We discarded the use of probability bounds in this work and determined protection levels as different percentages (e.g., 0.5, 1.0, and 1.5%) of the number of uncertain coefficients of each constraint. These numbers provided a good description of the tradeoff between robustness and optimality. For these three protection levels, we determined infeasibility rates by simulation as described in the following section.

Simulation Experiments

Because the probability bounds were not used to determine the protection levels, we had no insight into the infeasibility rates produced by the protection levels. We therefore estimated these rates by simulation experiments. We simulated the volume for each scenario (SCE10 and SCE15) and used this volume in the deterministic and robust solutions to evaluate the performance of the harvest decisions in satisfying demand constraints. In other words, the road decisions and the selection of stands to be harvested were fixed, and the harvest levels were recomputed with the simulated volume coefficients. Because no maximum capacity on arcs was assumed, the flow feasibility did not need to be assessed. The total simulated timber production was compared with the total minimum demand, and a simulation was considered infeasible if the production fell under the minimum requirement. Because specific flow details are not required to determine whether a simulation was feasible or not, the flow variable was not recomputed. A total of 1,000 simulations were performed for each scenario, and the solution of the deterministic model was compared with the solutions of the robust models with the three protection levels used (labeled as PL0.5, PL1.0, and PL1.5). In addition, we used the buffer estimated by robust models (i.e., the value of the protection function) in a deterministic framework; that is, the deterministic model was run with the modified demand levels given by the original demand plus the buffer. Therefore, we compared the effect of using a robust approach instead of a deterministic one with manually imposed buffers on the quality of the harvest and road construction decisions. For each simulated scenario, we evaluated the performance of the deterministic decisions (DET), three robust decisions (ROB), and three deterministic solutions with manually imposed buffers (BUF). The simulations were performed in MS Excel on an Intel Core 2 Quad 2.5 GHz with 4 GB of RAM.

Table 1. Numerical results of the optimization process.

Problem	No. of Constraints	No. of Binary variables	No. of Total variables	CPU time (min)	Residual gap (%)[a]
DET	4,025	2,532	5,017	60	2.3, 2.3
ROB_PL0.5	5,318	2,532	6,313	900	2.6, 2.7
ROB_PL1.0	5,318	2,532	6,313	900	2.7, 2.9
ROB_PL1.5	5,318	2,532	6,313	900	2.8, 3.1
BUF_PL0.5	4,025	2,532	5,017	60	2.5, 2.4
BUF_PL1.0	4,025	2,532	5,017	60	2.7, 2.5
BUF_PL1.5	4,025	2,532	5,017	60	2.4, 2.5

[a] Residual gap is shown for the two scenarios of volume estimate precision (SCE10, SCE15).

Results

Considering that solving this type of problems to optimality is virtually impossible, the models were run for a fixed period of time to obtain reasonably small gaps. The deterministic models were run for 1 hour and obtained gaps between 2.3 and 2.7%. The robust models were run for 15 hours to obtain similar gaps (except for SCE15 of the ROB_PL1.5, for which we obtained a 3.1% gap). We will refer to this greater difficulty in solving robust models in the Discussion section. The results of the optimization process are presented in Table 1.

The simulation experiments showed that the infeasibility rates of the deterministic decisions were considerably reduced as higher protection levels were used to find robust decisions (Figure 3). The infeasibility rates of approximately 49% observed in the deterministic decisions (protection level 0) dropped to 1.1 and 1.3% for scenarios SCE10 and SCE15, respectively, when the protection level was 1.5%. The decisions from the deterministic models with modified demand levels (BUF) also showed lower infeasibility rates, although these values were higher than those from the robust models (3.8% in SCE10 and 2.9% in SCE15). The need for higher harvest levels than those in the optimal deterministic solution caused a reduction in the objective function of the robust and buffered solutions (Figure 3). This reduction is small and consistently lower for the robust decisions than for the buffer solutions.

The robust models harvested a greater quantity of timber than the deterministic model because the former needed to guarantee the fulfillment of the minimum demand constraints for different values of the volume coefficients (Figure 4). As expected, the higher the protection level, the higher the timber production was. Furthermore, the higher the volume estimate error (Figure 4B), the higher the timber production required for an equivalent protection level was.

The road and harvest decisions from the robust models were different from those of the deterministic decisions. For the sake of brevity, only the decisions of scenario SCE10 are shown in Figure 5. The decisions of SCE15 were similar. In the first planning period, 61 (72) stands and 98 (112) segment roads were scheduled differently in SCE10 (SCE15) when the highest protection level was used to obtain the robust solutions (Figure 5A and C). The decisions of the buffer strategy were slightly different from those of the traditional deterministic model because modified demand levels were

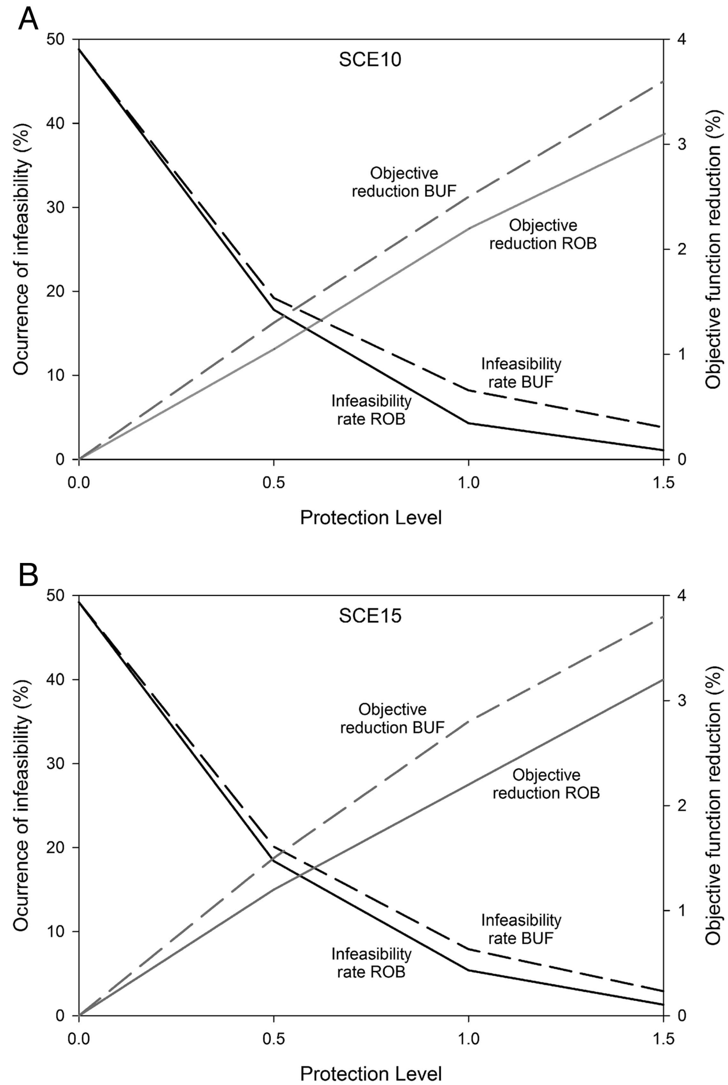

Figure 3. Reductions in the percentage of infeasible simulations and in the objective function were observed in both scenarios when higher protection levels were used. The buffer strategy showed higher infeasibility rates and higher losses in the objective function compared with the robust decisions. (A) SCE10. (B) SCE15.

used. In this case, 39 (44) stands and 67 (75) road segments were scheduled differently in SCE10 (SCE15). Although the buffer size was the same as that determined by the robust model, the decisions of the buffer strategy tended to follow a spatial pattern similar to that for the deterministic decisions (Figure 5A and B).

Because of the higher harvest levels required to obtain protection against infeasibility, both the robust and buffer models increased the number of stands and the area harvested throughout the planning horizon (Table 2). Surprisingly, in both scenarios the robust decisions generally required a smaller road network than the deterministic and buffer strategy models. The buffer model required even more roads than deterministic solutions. No major changes in the costs were observed among the different models. However, the ROB solutions consistently outperformed the BUF solutions. The road-building costs tended to be lower as more robustness was required.

Discussion

Our results suggest that using a robust optimization approach rather than a deterministic formulation to solve a road-building and harvest-scheduling problem produces solutions that are less sensitive to random errors in the volume estimates at the expense of a slight reduction in the objective function. Even in the case in which the appropriate buffer is used to modify the minimum demand levels of the deterministic models, the robust formulation produced more stable solutions and smaller losses in the objective function.

The robust optimization approach operates by determining a buffer or overproduction amount for each constraint for which feasibility is highly desirable. Because the ability to meet the minimum demand constraints is highly desirable in our case, the harvest levels of robust solutions were higher than those of the deterministic solutions. As expected, the greater the need for constraint fulfillment, the higher the level of timber production was. In addition, when the volume coefficients are more uncertain (i.e., scenario SCE15), an additional timber harvest is needed to guarantee a certain level of constraint satisfaction. This increase in the harvest level occurs because the wider range of possible values increases the magnitude of the eventual negative deviations against which we seek protection to satisfy the minimum demand levels. Although buffers can be used to manually modify the minimum demand levels, the robust approach estimates this buffer based on the decisions actually made as well as

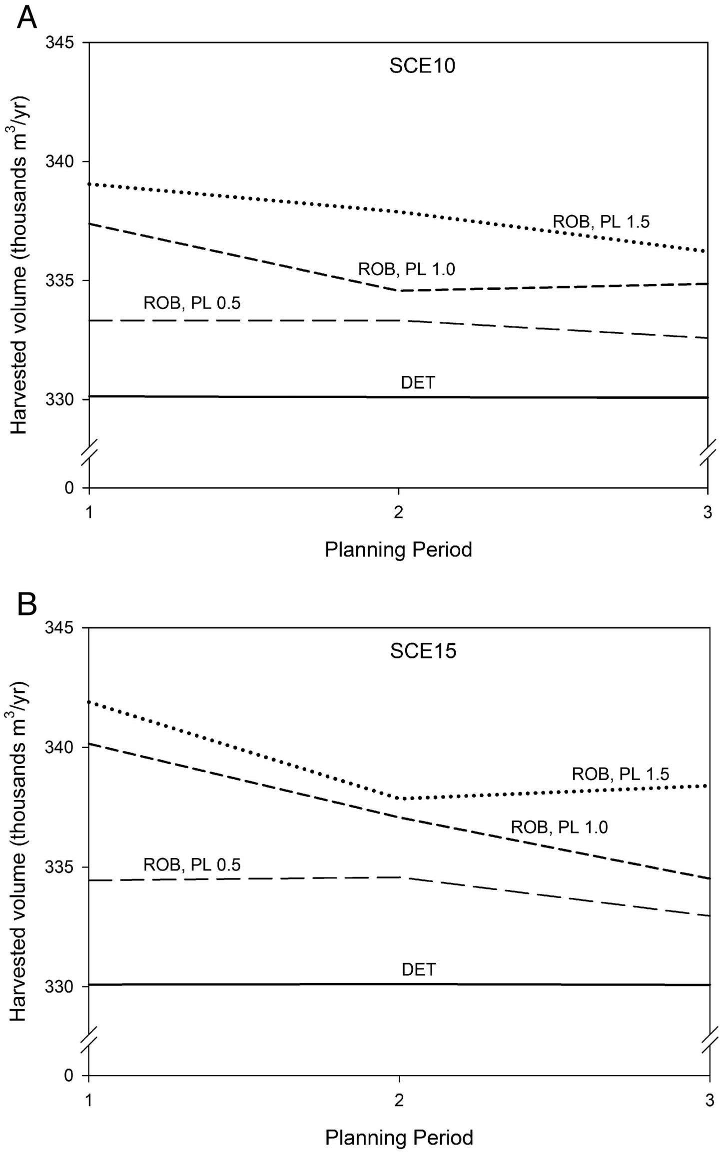

Figure 4. Harvest levels in the robust optimization models increased in both scenarios when a higher protection level (PL) was used. A lower timber harvest was required when the volume estimate error was smaller (SCE10). (A) SCE10. (B) SCE15.

the degree of data uncertainty, thus providing optimal "stable" decisions rather than the traditional optimal solutions. We demonstrated this improvement using a comparison between the robust approach and what we referred to as the buffer strategy. When the same buffer obtained by the robust models was used to modify the demand levels of the deterministic model, the robust solutions outperformed the deterministic decisions both in terms of infeasibility rates and objective function value. Although one might conclude that the buffer strategy performed almost as well as the robust approach, we note that finding the buffer levels by means other than the approach presented in this work is quite difficult. The amount of buffer obtained by the robust models was "optimal" within the range of the solution gap, and, therefore, the use of any other buffer level would probably produce inferior results. Because it is unlikely that integer models will be solved to optimality, the presence of a gap in the solutions precludes an exact comparison of their quality. Obtaining the same gap for different models to standardize the comparison is also difficult because the gaps progress by discrete steps rather than by continuous movements. However, the greater gap observed in the robust models might represent additional room for improvement of their solutions and may therefore enhance the benefits of the robust formulations.

The volume estimate error affected the management decisions. The robust models scheduled both the harvest and road construction in a manner different from that of the deterministic model to take advantage of the uncertain timber yield throughout the planning horizon in a cost-efficient way. Although the robust models and the buffer strategy were set up with the same minimum level of production, their decisions differed to a large extend. Whereas the buffer strategy tended to follow a spatial pattern similar to that for deterministic decisions (with the extra requirement to harvest additional timber), the robust approach favored the spatial concentration of the harvest operations in each period. Although clustering operations may sound counterintuitive if spatial diversification is desired, they reduced the cost of road construction (Table 2) and therefore balanced the tradeoff between the optimality and the robustness of solutions. We note that although the decisions are scenario-specific (i.e., influenced by stand volume, roads network, and

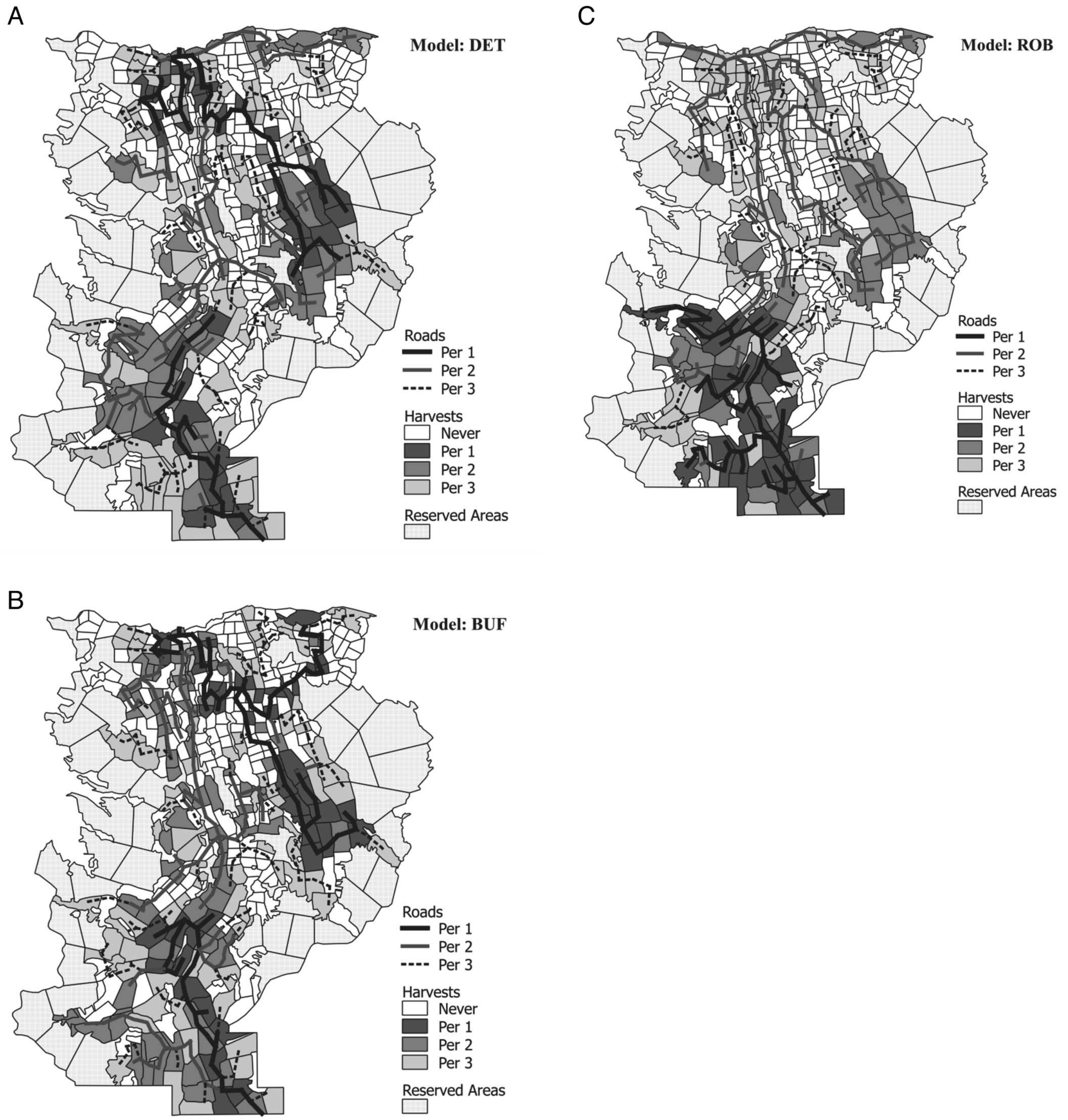

Figure 5. Harvesting and road-building decisions were different among the deterministic (DET [A] and BUF [B]) and robust models (ROB [C]). The buffer strategy tended to produce a decision pattern similar to that of the deterministic model.

demands), we still expect that robust optimization will yield better protection and better solutions than the deterministic and buffer strategies. We used short planning periods, and, in these cases, a solution is more likely to be implemented than with long planning periods. Hence, the robust solutions offer better recourse when it comes time to replan for the next period(s). From this perspective, we believe that robust optimization has good potential for this type of planning problems. For larger problems, the increased availability of stands and arcs for potential roads will probably provide more flexibility for finding robust solutions, therefore enhancing its benefits over a determinist approach.

The robust approach could also be applied to explore the value of the inventory precision or the value of the information. If we consider the deterministic formulation of the problem as the perfect information scenario, the decrease in the value of the objective function of a robust model for a given estimate error may represent the cost of not carrying a perfect inventory. Reducing the estimate error of the inventory will translate into a better and more robust objective function value, although it may still be worse than the deterministic solution. The improvement between the two robust objective values would represent the value of the improvement in the inventory precision.

Table 2. Summary of harvest results for the planning models (SCE10/SCE15).

Model	No. stands	Total area (ha)	Roads (km)	Average cost ($/m^3)		
				Harvest	Transportation	Roads
Det	218	4,804	108.7	2.85	1.77	2.55
Rob0.5	223/223	4,827/4,835	110.2/108.7	2.89/2.87	1.72/1.71	2.62/2.55
Rob1.0	224/226	4,911/4,885	106.7/110.0	2.86/2.88	1.69/1.68	2.50/2.62
Rob1.5	226/231	4,923/4,995	107.6/105.8	2.81/3.01	1.75/1.77	2.48/2.45
Buf0.5	223/219	4,835/4,821	112.9/110.9	2.88/2.85	1.82/1.78	2.59/2.58
Buf1.0	223/227	4,897/4,903	119.5/110.1	2.85/2.91	1.82/1.79	2.74/2.57
Buf1.5	224/228	4,868/4,976	113.1/116.0	2.90/2.88	1.76/1.77	2.59/2.62

For each constraint with uncertain coefficients and for which feasibility is desirable, a new term is added. This term, known as the protection function, represents an optimization problem in itself. For each solution vector, the process looks for the buffer required to meet the constraint for a given user-defined protection level. From a mathematical point of view, the difference between the robust and deterministic (including the buffer approach) formulations is that the former imposes bounds on the variables associated with the uncertain coefficients. From Equation 21, it is clear that harvest decision Y_{st} is upper-bounded by $(z_t + p_{st})/\hat{v}_{st}$, which implies that the higher uncertainty forces larger values for z and p, and therefore a larger buffer if the corresponding stand is harvested. Decisions that contribute to both a higher objective value and small buffers are preferred, thereby providing high-quality solutions. Although the same protection levels were used in our analysis for all constraints, these levels can differ to express different degrees of importance. For example, protection levels can emphasize the importance of the first period over the remainder of the planning horizon or the differences among products if a multiproduct model is used. In this case, the manual estimation of the proper buffers for each product becomes more difficult, and the robust optimization approach becomes more useful. The importance of demand centers can be handled with the parameter δ_{nt}, which represents the fraction of the buffer to be sent to each destination point. We consider that our initial setting for the total demand (70% of the available volume) will not change the effect of using the robust optimization approach. However, the higher the proportion of the volume to be harvested, the higher the chance of obtaining infeasible robust solutions because extra harvests are required.

The robust models contain more variables and constraints than the original models. In our case, because the harvested volume in each period t comes from the potential harvest of any of the S stands, $S \times T$ ($431 \times 3 = 1{,}293$) new constraints must be added (given that all stands are available and all have uncertain volume). For each of these constraints, as well as for each period, a new variable must also be created, that is, $S \times T + T$ ($1{,}293 + 3 = 1{,}296$). Although the increase in the model size does not appear to be particularly important, the robust models were harder to solve than the deterministic model because additional protection levels were used. Despite the extended solution time used for the robust models, the residual gaps were greater than those for the deterministic models. This result suggests that the robust formulations are weaker than the deterministic formulations and that extra effort should be made to obtain the same gaps as those for the deterministic approach (i.e., increased solution times and other strengthening and solution techniques).

Our approach includes some assumptions. For example, the uncertain coefficients must be noncorrelated in the same row, which means that the volume coefficients of each constraint should be independently distributed. This situation forced us to assume that the stand volume uncertainties originate from random and spatially independent errors, which is usually expected if a recent cruise of the area is used. However, uncertain trends in data and catastrophic events such as fires and pest attacks cannot be properly modeled with the current assumptions, and, hence, further research is needed to address this limitation. It is important to note, however, that a certain degree of independency must be assumed if we are to seek stable solutions to address uncertainty. If decisions are highly correlated, then we cannot take advantage of diversification strategies and only the worst-case scenario solutions would become relevant.

The uncertain values are also assumed to be uniformly and symmetrically distributed, which can lead to the loss of information if the uncertainties are well described by a different probability distribution. This situation could be handled by considering a uniform range that embraces, for instance, 99% of the observations of the original distribution of uncertain values (Palma and Nelson 2009). Although more conservative solutions will be obtained than when the true distribution is used, this simplification assists in assuring that the robust models are easier to solve. Another disadvantage of the methodology is the impossibility of accurately determining the protection levels required to deliver specific infeasibility rates. Although probability bounds can be used to estimate them (Bertsimas and Sim 2004), they are loose and produce more conservative decisions than desired (Palma and Nelson 2009). As in this work, simulation experiments can be used to evaluate the infeasibility rates.

Further research should address the previously mentioned limitations. Possible methods for inclusion of some levels of independency as well as the asymmetry of the uncertain coefficients have been proposed recently (Chen et al. 2007), and their applicability should be explored. In addition, more accurate probability bounds to relate the probability of constraint violation and protection levels would improve this methodology by eliminating the need for simulation experiments. Finally, the application of alternative solution techniques (i.e., Lagrangean relaxation and heuristics) to solve robust road-building formulations would also be of interest in solving larger problems quickly and with reduced gaps.

Conclusions

We presented a robust formulation of a road-building and harvest-scheduling problem with random volume coefficients for situations in which the feasibility of minimum demand constraints is highly desired, although the same approach can be used to address other constraints. The decisions obtained with this formulation were less sensitive to volume uncertainties and more efficient in terms of the objective value than those of the traditional formulations or buffering strategies used. The appropriate amount of buffer required for each constraint was determined in conjunction with the decision variables in such a way that the tradeoff between cost and robustness was explicitly considered in the formulation. Although larger models were obtained, the approach remains the model type (i.e., linear mixed integer), and commercial integer optimization software can be used to solve the problem. However, larger gaps were also observed, suggesting that additional strengthening or different solution approaches may be necessary. The formulation allows for different levels of protection (robustness) against constraint infeasibility, which make it possible to evaluate the tradeoff between robustness and optimality. Because the current probability bounds that relate protection levels and feasibility rates represent only weak estimates of the probability of constraint satisfaction, simulation experiments were performed to obtain the infeasibility rates. The independency assumption among uncertain coefficients appears to be the main limitation of the methodology if spatial correlation in errors must be considered, suggesting that further research is needed in this area.

Literature Cited

ALONSO-AYUSO, A., E.F. ESCUDERO, M. GUIGNARD, M. QUINTEROS, AND A. WEINTRAUB. 2009. Forestry management under uncertainty. *Ann. Oper. Res.* 190(1):17–39.

ANDALAFT, N., P. ANDALAFT, M. GUIGNARD, A. MAGENDZO, A. WAINER, AND A. WEINTRAUB. 2003. A problem of forest harvesting and road building solved through model strengthening and Lagrangean relaxation. *Oper. Res.* 51:613–628.

BARE, B.B., AND G.A. MENDOZA. 1992. Timber harvest scheduling in a fuzzy decision environment. *Can. J. For. Res.* 22:423–428.

BEN-TAL, A., AND A. NEMIROVSKI. 2002. Robust optimization—Methodology and applications. *Math. Program.* 92:453–480.

BEN-TAL, A., AND A. NEMIROVSKI. 2000. Robust solutions of linear programming problems contaminated with uncertain data. *Math. Program.* 88:411–424.

BERTSIMAS, D., AND A. THIELE. 2006. A robust optimization approach to inventory theory. *Oper. Res.* 54:150–168.

BERTSIMAS, D., AND M. SIM. 2004. The price of robustness. *Oper. Res.* 52:35–53.

BERTSIMAS, D., AND M. SIM. 2003. Robust discrete optimization and network flows. *Math. Program.* 98:49–71.

BERTSIMAS, D., D. PACHAMANOVA, AND M. SIM. 2004. Robust linear optimization under general norms. *Oper. Res. Lett.* 32:510–516.

BEVERS, M. 2007. A chance constraint estimation approach to optimizing resource management under uncertainty. *Can. J. For. Res.* 37:2270–2280.

BIRGE, J.R., AND F. LOUVEAUX. 1997. *Introduction to stochastic programming*, 1st ed. Springer, New York. 421 p.

BOYCHUK, D., AND D.L. MARTELL. 1996. A multistage stochastic programming model for sustainable forest-level timber supply under risk of fire. *For. Sci.* 42:10–26.

BOYLAND, M., J. NELSON, F.L. BUNNELL, AND R.G. D'EON. 2006. An application of fuzzy set theory for seral-class constraints in forest planning models. *For. Ecol. Manage*. 223:395–402.

CHEN, X., M. SIM, AND P. SUN. 2007. A robust optimization perspective on stochastic programming. *Oper. Res.* 55:1058–1071.

CHUNG, G., K. LANSEY, AND G. BAYRAKSAN. 2009. Reliable water supply system design under uncertainty. *Environ. Model. Softw.* 24(4):449–462.

CLARK, M.M., R.D. MELLER, AND T.P. MCDONALD. 2000. A three-stage heuristic for harvest scheduling with access road network development. *For Sci.* 46:204–218.

DAVIS, L., AND L. JOHNSON. 1987. *Forest management*, 3rd ed. McGraw Hill, New York. 690 p.

EL GHAOUI, L., F. OUSTRY, AND H. LEBRET. 1998. Robust solutions to uncertain semidefinite programs. *Siam J. Optimiz.* 9:33–52.

ELLS, A., E. BULTE, AND G.C. VANKOOTEN. 1997. Uncertainty and forest land use allocation in British Columbia: Vague priorities and imprecise coefficients. *For. Sci.* 43:509–520.

ERIKSSON, L.O. 2006. Planning under uncertainty at the forest level: A systems approach. *Scand. J. For. Res.* 21:111–117.

GASSMANN, H.I. 1989. Optimal harvest of a forest in the presence of uncertainty. *Can. J. For. Res.* 19:1267–1274.

GUIGNARD, M., C. RYU, AND K. SPIELBERG. 1998. Model tightening for integrated timber harvest and transportation planning. *Eur. J. Oper. Res.* 111:448–460.

HOF, J.G., B.M. KENT, AND J.B. PICKENS. 1992. Chance constraints and chance maximization with random yield coefficients in renewable resource optimization. *For. Sci.* 38:305–323.

HOF, J.G., J.B. PICKENS, AND E.T. BARTLETT. 1986. A maxmin approach to nondeclining yield timber harvest scheduling problems. *For. Sci.* 32:653–666.

HOF, J.G., K.S. ROBINSON, AND D.R. BETTERS. 1988. Optimization with expected values of random yield coefficients in renewable resource linear-programs. *For. Sci.* 34:634–646.

HOGANSON, H.M., AND D.W. ROSE. 1987. A model for recognizing forestwide risk in timber management scheduling. *For. Sci.* 33:268-282.

KALL, P., AND S.W. WALLACE. 1994. *Stochastic programming*, 2nd ed. John Wiley and Sons, Chichester, UK. 307 p.

MANESS, T., AND R. FARRELL. 2004. A multi-objective scenario evaluation model for sustainable forest management using criteria and indicators. *Can. J. For. Res.* 34:2004–2017.

MARSHALL, P.L. 1987. Sources of uncertainty in timber supply projections. *For. Chron.* 63:165–168.

MENDOZA, G.A., B.B. BARE, AND Z.H. ZHOU. 1993. A fuzzy multiple objective linear-programming approach to forest planning under uncertainty. *Agr. Syst.* 41:257–274.

MOWRER, H.T. 2000. Uncertainty in natural resource decision support systems: Sources, interpretation, and importance. *Comput. Electron. Agric.* 27:139–154.

MULVEY, J.M., R.J. VANDERBEI, AND S.A. ZENIOS. 1995. Robust optimization of large-scale systems. *Oper. Res.* 43:264–281.

OLSSON, L. 2007. Optimal upgrading of forest road networks: Scenario analysis vs. stochastic modelling. *For. Policy Econ.* 9:1071–1078.

ORDONEZ, F., AND J.M. ZHAO. 2007. Robust capacity expansion of network flows. *Networks* 50:136–145.

PALMA, C.D., AND J.D. NELSON. 2009. A robust optimization approach protected harvest scheduling decisions against uncertainty. *Can. J. For. Res.* 39:342–355.

PALMA, C.D., AND J.D. NELSON. 2010. Bi-objective multi-period planning with uncertain weights: A robust optimization approach. *Eur. J. For. Res.* 129:1081–1091.

PICKENS, J.B., AND P.E. DRESS. 1988. Use of stochastic production coefficients in linear-programming models—Objective function-distribution, feasibility, and dual activities. *For. Sci.* 34:574–591.

PICKENS, J.B., AND J.G. HOF. 1991. Fuzzy goal programming in forestry—An application with special solution problems. *Fuzzy Set Syst.* 39:239–246.

PICKENS, J.B., J.G. HOF, AND B.M. KENT. 1991. Use of chance-constrained programming to account for stochastic variation in the A-matrix of large-scale linear programs: A forestry application. *Ann. Oper. Res.* 31:511–526.

REGAN, H.M., M. COLYVAN, AND M.A. BURGMAN. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecol. Appl.* 12:618–628.

RICHARDS, E.W., AND E.A. GUNN. 2000. A model and tabu search method to optimize stand harvest and road construction schedules. *For. Sci.* 46:188–203.

STIRN, L.Z. 2006. Integrating the fuzzy analytic hierarchy process with dynamic programming approach for determining the optimal forest management decisions. *Ecol. Model.* 194:296–305.

WEINTRAUB, A., AND A. ABRAMOVICH. 1995. Analysis of uncertainty of future timber yields in forest management. *For. Sci.* 41:217–234.

WEINTRAUB, A., G. JONES, M. MEACHAM, A. MAGENDZO, A. MAGENDZO, AND D. MALCHUK. 1995. Heuristic procedures for solving mixed-integer harvest scheduling-transportation-planning models. *Can. J. For. Res.* 25:1618–1626.

WEINTRAUB, A., AND D. NAVON. 1976. Forest management planning model integrating silvicultural and transportation activities. *Manage. Sci.* 22:1299–1309.

WEINTRAUB, A., AND J. VERA. 1991. A cutting plane approach for chance constrained linear-programs. *Oper. Res.* 39:776–785.

ZIMMERMANN, H.J. 1996. *Fuzzy set theory and its applications*, 3rd ed. Kluwer Academic, Boston, MA. 435 p.