Intl. Trans. in Op. Res. 10 (2003) 413–431

INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

Short-term harvest planning including scheduling of harvest crews

J. Karlsson[a], M. Rönnqvist[a,b] and J. Bergström[b]

[a]*Division of Optimization, Linköping Institute of Technology, SE-58183 Linköping, Sweden,* [b]*The Forest Research Institute of Sweden, Uppsala Science Park, SE-751 83 Uppsala, Sweden*

Received 4 March 2002; received in revised form 4 November 2002; accepted 4 February 2003

Abstract

The problem we consider is short-term harvesting planning for a total planning period of 4–6 weeks where we want to decide the harvest sequences or schedules for harvest crews. A schedule is an order or sequence of harvest areas assigned to each crew. The harvesting of areas is planned in order to meet industrial demand. The total cost includes harvesting, transportation, and storage. One considerable cost is due to the quality reduction of logs stored at harvest areas. There are a number of restrictions to be considered. Areas are of varying size and the composition of assortments in each area is different. Each harvest team has different skills, a different home base, and different production capacity. Another aspect is the road network. There is a cost related to road opening (restoring, snow removal). In this paper, we develop a mixed integer programming (MIP) model for the problem. The schedules are represented by 0/1 variables. With a limited number of schedules, the problem can be solved by a commercial MIP solver. We have also developed a heuristic solution approach that provides high-quality integer solutions within a distinct time limit to be used when more schedules are used. Computational results from a major Swedish forest company are presented.

Keywords: forestry; operations research; harvesting; scheduling; heuristics; integer programming; modeling.

Introduction

At Swedish forest companies the wood flow planning, including harvest and transportation planning, occurs in different stages. An overview of wood flow problems in Swedish forestry is given in Carlsson and Rönnqvist (1998). Larger forest companies in Sweden have an organization whereby operations are divided into smaller regions, which may be composed of one or several districts. In each district, both medium and short-term harvesting plans are made. On a medium or annual level, planning is based on several-year-long harvesting plans and is aimed to identify which harvest areas to harvest in given months in order to balance supply and industrial demand.

Here it is important to consider accessibility during different parts of the year. On a shorter time horizon, plans are often made for the next few weeks. Plans at different levels are continuously updated in rolling planning horizons.

The harvest planning at a district includes decisions about which areas to harvest and which crews should do it. The selection of harvesting areas strongly affect the production level of different assortments. Possibilities of varying the production level are described in Brunberg (2001). This involves, for example, selecting areas with larger trees, areas with particular assortment mix, allowing overtime, changing or adding equipment for crews. The planning also includes transportation planning, road maintenance decisions, and control of storage in the forest and at terminals. The weather conditions affect annual harvest planning. An annual plan considers accessibility of both roads and harvest areas, and makes sure that monthly demand of each assortment or species can be harvested. In Karlsson, Rönnqvist and Bergström (2002), a mathematical model for the annual problem has been developed and tested.

The problem we consider in this paper is an operational planning problem at district level. The total planning period is typically five to six weeks. The main decisions deal with the same aspects as the annual planning problem, but there are differences. In operational planning, each crew is given a schedule, which means a defined sequence of areas to be harvested during the planning period. Here, costs for moving equipment are considered as the actual sequence of areas is known. In annual planning, the sequence of harvest areas to be harvested during a month is not defined. Furthermore, it is important to take into consideration the storage level and age of logs, as a significant cost is due to the quality reduction of products stored. The short-term planning problem, which in this case has a planning horizon of 5–6 weeks, starts from a list of areas suitable for the particular planning period. The planning is done on a rolling horizon and 3–4 weeks is enough to consider actual situations at industries and for crews. However, to take away end effects, a number of weeks is added. The total supplies of assortments in harvest areas must ensure that the industrial demand is satisfied. Moreover, the harvesting areas must be accessible. This list may be a result of an annual plan or may be given by a planner at a given time. This means that the accessibility of roads and areas is considered as given at this level of planning.

A decision support system concerning short-term harvesting planning, used in the Chilean forestry industries is called OPTICORT (Epstein et al., 1999a, 1999b). This system is based on a linear model with a column generation procedure. The main decisions are related to which stands to harvest, what type of machinery to use for each operation, what volume to cut each week or period, and what products should be delivered to different destinations to satisfy demand. Decisions concerning which bucking patterns to use are included. The total planning period is three months. There are several models supporting decisions about scheduling equipment on short time horizons. Jarmer and Sessions (1992) describe LoggerPC, a program that provides a physical feasible design analysis for harvesting by cable. The PLANEX system (Epstein et al., 1999a, 1999b) determines the optimal location of towers, roads, and skidder operations.

Tactical and operational harvest planning often includes discrete decisions about specific areas or roads, which leads to MIP problems. The presence of integer variables and the size of real-world problems often requires heuristic methods to achieve practical solution times. A long-term spatially constrained harvest scheduling problem is studied in Yoshimoto, Brodie and Session (1994). They have developed a heuristic procedure in which the problem is partitioned into numerous subproblems. Weintraub et al. (1995) describe a heuristic algorithm for solving a tactical problem for two to three decades including harvesting decisions and road building. The solution procedure iterates between solving relaxed LP versions of the model and applying rules to fix fractional variables to integer variables. Nelson and Brodie (1990) give a comparison of a heuristic with an optimization method applied on a three-decade tactical harvesting problem.

Our short-term harvesting problem has to be compatible with the annual planning. This is achieved through a hierarchical approach where tactical and operational harvesting plans are made by two different models. The output from the tactical gives input, i.e., a list of possible areas, to the operational model. In practice, plans at different levels are continuously updated. A surplus of possible harvesting areas is used as a starting point for the operational planning. To achieve compatible decision-making on different levels, different approaches have been used. Hierarchical methods have been developed to integrate tactical and strategic planning (Weintraub and Cholaky, 1991). McNaughton, Rönnqvist, and Ryan (2000) present an integrated model, which incorporates both strategic long-term goals and detailed area-sensitive tactical planning.

We start by developing a mathematical model. Furthermore we are interested to find good solutions in short times. This often leads to heuristic procedures. The main decisions in our problem are similar to the main decisions in OPTICORT, except for estimating bucking patterns. We do not consider bucking patterns, as Swedish harvesters are equipped with on-board computers that provide cutting instructions. We also include the scheduling of crews to harvest areas. Another important aspect is storing at areas and terminals. Considerable cost is related to storing outdoors, because of quality decrease. The main new aspects in our problem are:

- scheduling of the harvesting crew;
- age-related storing costs.

There are a large number of possible harvesting schedules for the teams, which correspond to a large number of binary variables. Furthermore, the consideration of different age classes leads to a large number of continuous variables and constraints in this problem.

Currently, planning is carried out manually. There is most often a limited number of qualified and experienced persons to make the short-term harvesting plans and they spend a large amount of time developing practical plans. At the same time, it is difficult to consider all aspects of the planning as it covers many components. Coordination between districts is very limited and it is only considered particular situations during the year.

We have developed an MIP model for the short-term harvesting planning problem. With a limited number of schedules for the harvesting crews it is possible to solve the problem to optimality directly by using a commercial integer solver. We have also developed a heuristic solution procedure that finds good solutions fast. It is not a production system, but it is intended to become an important part of a future decision support tool.

The remainder of this paper is organized as follows. In the next section the short-term harvesting planning problem is described. Next we present the mathematical formulation of the problem. Then the case study is described, and the solution methods and the numerical results are given. Finally some concluding remarks and suggestions for further work are given.

Problem formulation

Harvest areas

Each harvest area is unique, with its own properties. It varies in size and in available volumes of assortments after harvesting. Areas require between 1 and 20 days work to harvest. The operations are either final felling (all trees in the area are harvested) or thinning (a proportion of the trees are harvested in order to improve the overall growth). Once an area is harvested it produces a given amount of assortments. The production time depends on the average size of the trees, because of equipment constraints. Therefore, the amount of different assortments harvested in a particular week for a crew depends on which areas are harvested. Each area is also connected to a particular road or road group. Table 1 gives some typical data for a number of areas in the case study.

Storage

An important aspect is the loss or reduction of value due to quality deterioration. This happens when logs are stored outdoors, at areas or terminals. The quality decrease is negligible during the cold season of the year. In the north of Sweden this usually corresponds to December through March. During the rest of the year this quality loss causes considerable costs. Spruce pulp (GM) is the most sensitive assortment and it can lose almost its total value after five weeks of storing. Pulp mills do not accept spruce pulp stored for more than four to five weeks. Table 2 gives an example of the value decrease given in SEK for different assortments and different storage times. In order to handle storage, we need to keep track of the age of the wood.

Harvest teams

A harvest crew or team consists of a harvester and a forwarder, and two working groups with two people in each. The harvester fells and bucks the trees. The forwarder collects all the piles on the area and moves the logs to pick-up points adjacent to the forest road. Each

Table 1
Illustration of information given for harvest areas

Name	Hours	Size	TT	GT	TM	GM	LM
Koffsan	20.8	0.237	217	—	100	—	—
Stormon	30.9	0.447	12	37	22	32	3
Dalbrants	29.8	0.419	16	252	46	152	73
Harte	18.2	0.195	45	13	22	15	—
Lunnsjo	24.3	0.296	83	407	44	207	—

Notes: The column 'Hours' gives the number of standard hours to harvest the area. Each crew then has a relative factor or efficiency rate against this. 'Size' is the average tree diameter given in metres: This may decide if certain crews can harvest it or not depending on their equipment. The quantities of the different assortments given in cubic meters are also given in the table. The different assortments are 'TT': Pine-timber, 'GT': Spruce-timber, and 'TM': Pine-pulp, 'GM': Spruce-pulp, and 'LM': Birch-pulp.

Table 2
Example of value decrease per m^3 for different assortments and age in SEK

Sort	Week 1	Week 2	Weeks 3	Week 4	Week 5	Week 6
TT	6.67	13.33	23.25	31	38.75	46.5
GT	6.67	13.33	13.25	31	38.75	46.5
LM	1.67	3.34	5	29.76	54.34	79
GM	0	20	20	281	476	505
TM	1.67	3.34	5	34	63	92

team has a unique capacity and efficiency depending on their equipment. Some teams concentrate on final felling, and others on thinning. Some teams are not able to harvest particular areas due to their equipment. For each combination of harvest area and team, there is a given production capacity and cost for harvesting and forwarding. The personnel in each harvest team have given home bases, and there is a specific travel distance and cost for each combination of working team and harvest area. Each team travels back and forth each working day.

Harvesting schedules

The main decision in the planning problem is to give each working team a schedule. A schedule is a sequence of areas, defining which areas and in what order these are to be harvested during the planning period. The length (time) of a schedule covers (at least) the number of working days during the planning period. For each planning period, there are a number of areas, which are potential areas to harvest. This correspond to about 1.5–2 times the actual demand. There are many possible schedules for each team. We generate a limited number of schedules, typically 1000–2000, for each team starting from their particular list of possible harvesting areas. Areas corresponding to short travel (of crew) and moving (of equipment) distances are more likely to be included. The procedure for generating sequences is described in the next section. Each schedule has a given cost. This, together with volumes harvested of different assortments in each area during each time period, and road conditions associated with each sequence, is computed. We note that each sequence consists of many pieces of information. In the model, each schedule will correspond to the contents of one column (or variable).

In Fig. 1 we illustrate a schedule, and the volumes of assortments harvested in each time period. The sequence starts with harvesting area 3, then areas 4, 7, and finally area 2. In this example, the planning period consists of five time periods. Each time period corresponds to five working days. Assume that area 3 requires seven days of harvesting. This means that 5/7 of the total volumes harvested in area 3 is produced in time period 1 and 2/7 in time period 2. Furthermore, area 4 requires five working days and the harvested volume is produced in time periods 2 and 3 (3/5 and 2/5 respectively). Area 7 is smaller and requires three days and is finished in time period 3. Finally, area 2 requires 13 days. The overall schedule stretches outside the planning period.

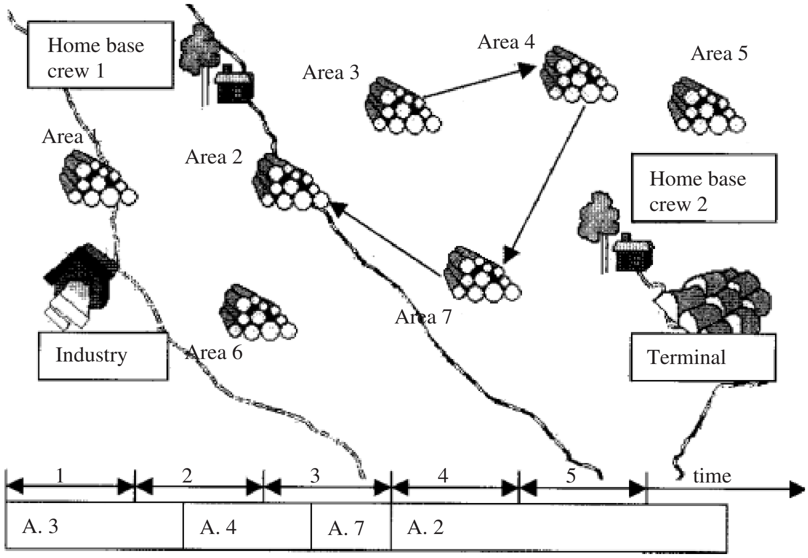

Fig. 1. Example of a harvesting sequence.

Industrial demand and storage

The industrial demand for each industry (saw, pulp- or paper mills) is given per week and assortment. Each industry has a given storage capacity with a given cost. The transportation cost depends on assortment and distance, and is proportional to the volume transported.

Roads

Each forest road is included in a group of forest roads. Each group of forest roads is assigned a specific ID number. Some areas are connected directly to the state road network. Most areas are connected to a forest road group, which in turn is connected to the state road network. The use of road group is standard, as single roads are regarded as too detailed. Also, there are areas that are connected to forest roads, which are connected to another forest road group before the state road network is reached. This hierarchy decides how maintenance and snow removal must be carried out. Accessibility of the state roads is guaranteed. The short-term plan is based on the assumption that all roads connected to the areas in the list are accessible to each road given a cost. Some forest roads need restoring to be accessible, e.g. in wintertime snow must be removed. The road dependence is illustrated in Fig. 2. In this example areas 1, 2, 3 and 4 are connected to road group 1, which is connected to the state road network. Areas 5, 6, and 7 are connected to road group 2, and areas 8, 9, and 10 to road group 3. Road group 3 is connected to road group 1 or 4 and road group 2 is connected to road group 1 or 5, before reaching the state road network. The road network can have any structure.

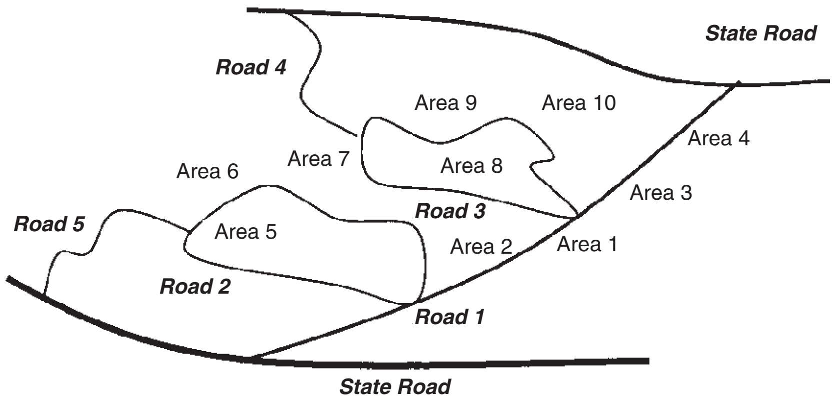

Fig. 2. Illustration of a possible road network.

Modelling

In this section, the mathematical model is presented. A premise for the model formulation is that a number of harvesting sequences is available for each team. The procedure for generating sequences is also given in this section.

Mathematical formulation

First the defined sets used in the model are introduced.

S: The set of all harvesting sequences for all working teams.
$SQ(q)$: The set of schedules that correspond to crew q.
$SA(i)$: The set of schedules in which area i is included.
I: The set of all harvesting areas suitable for harvest during the planning period and areas with stocks harvested before the planning period started.
J: The set of industries with demands from the district.
P: The set of products at harvesting areas.
L: The set of terminals it is possible to use for storing.
T: The number of time periods during the total planning period.
A: The different age classes of the harvested timber, i.e., the number of time periods it has been stored (given in weeks).
R: The set of roads connecting the areas to public roads.
$R^P(r)$: Roads adjacent to road r, i.e., roads connecting r to a higher level in the road hierarchy.
$R(i)$: The set of roads r that connects area i to the next level of roads (or public road).

The data needed to formulate the model is given below. All costs are unit costs when applicable.

K_s^S = cost of schedule s including harvesting, forwarding, travelling, and moving equipment.
K_{rt}^R = cost for using road r during time period t.
P_{ijk}^U = unit transportation cost from area i to industry j of assortment k.

P^V_{ijk} = unit transportation cost from area i to terminal j of assortment k.
P^W_{ijk} = unit cost for transportation from terminal l to industry j of assortment k.
R^R_{ka} = value reduction in areas/terminals of assortment k of age a.
R^I_{ka} = storage cost at industries of assortment k of age a.
a^A = the highest age class considered, corresponding to timber stored a^Aweeks or more.
S_i = total supply at area i.
d_{jkt} = demand at the industry j of assortment k in time period t.
b_{jkt} = storage capacity of product k at industry j during time period t.
a_{ikts} = amount of assortment k harvested in area i in period t using schedule s.
h_{sit} = 1 if (part of) area i is harvested during time period t in schedule s, 0 otherwise.

The variables used in the model are as follows.

$$x_s = \begin{cases} 1 \text{ if sequence } s \text{ is used,} \\ 0 \text{ otherwise.} \end{cases}$$

$$y_{rt} = \begin{cases} 1 \text{ if road } r \text{ is used, during time period } t, \\ 0 \text{ otherwise.} \end{cases}$$

u_{ijkat} = flow from area i to industry j of assortment k of age a during period t.
v_{ilkat} = flow from area i to terminal l of assortment k of age a during period t.
w_{ljkat} = flow from terminal l to industry j of assortment k of age a during period t.
z_{jkat} = amount of assortment k of age a used at industry j during period t.
L^F_{ikat} = storage in area i of assortment k of age a during period t.
L^I_{lkat} = storage at terminal l of assortment k of age a during period t.
L^I_{jkat} = storage at industry j of assortment k of age a during period t.

Now the model can be formulated as:

$$\begin{aligned} \min & \sum_{s\in S} K^S_s x_s + \sum_{r\in R}\sum_{t\in T} K^R_{rt} y_{rt} \\ & + \sum_{i\in I}\sum_{j\in J}\sum_{k\in P}\sum_{a\in A}\sum_{t\in T} P^U_{ijk} u_{ijkat} + \sum_{i\in I}\sum_{l\in L}\sum_{k\in P}\sum_{a\in A}\sum_{t\in T} P^V_{ilk} v_{ilkat} + \sum_{l\in L}\sum_{j\in J}\sum_{k\in P}\sum_{a\in A}\sum_{t\in T} P^W_{ijk} w_{ljkat} \\ & + \sum_{i\in I}\sum_{k\in P}\sum_{a\in A}\sum_{t\in T} R^R_{ka} L^F_{ikat} + \sum_{l\in L}\sum_{k\in P}\sum_{a\in A}\sum_{t\in T} R^R_{ka} L^T_{lkat} + \sum_{j\in J}\sum_{k\in P}\sum_{a\in A}\sum_{t\in T} R^I_{ka} L^I_{jkat} \end{aligned}$$

subject to:

$$\sum_{s\in SA(i)} x_s \leq 1 \quad \forall i \tag{3.1}$$

$$\sum_{s\in SQ(q)} x_s = 1 \quad \forall q \tag{3.2}$$

$$L^F_{ikat} = \sum_{s \in SA(i)} a_{ikts} x_s - \sum_{j \in J} u_{ijkat} - \sum_{l \in L} v_{ilkat} \qquad \forall i, k, t, a = 1 \tag{3.3a}$$

$$L^F_{ikat} = L^F_{ik,a-1,t-1} - \sum_{j \in J} u_{ijkat} - \sum_{l \in L} v_{ilkat} \qquad \forall i, k, t, 2 \leqslant a < a^A \tag{3.3b}$$

$$L^F_{ikat} = L^F_{ika-1,t-1} + L^F_{ika,t-1} - \sum_{j \in J} u_{ijkat} - \sum_{l \in L} v_{ilkat} \qquad \forall i, k, t, a = a^A \tag{3.3c}$$

$$L^I_{lkat} = \sum_{i \in I} v_{ilkat} - \sum_{j \in J} w_{ijkat} \qquad \forall l, k, t, a = 1 \tag{3.4a}$$

$$L^I_{lkat} = L^I_{lk,a-1,t-1} - \sum_{i \in I} v_{ilkat} - \sum_{j \in J} w_{ilkat} \qquad \forall l, t, 2 \leqslant a < a^A \tag{3.4b}$$

$$L^I_{lkat} = L^I_{lk,a-1,t-1} + L^L_{lka,t-1} + \sum_{i \in I} v_{ilkat} - \sum_{j \in J} w_{ljkat} \qquad \forall l, k, t, a = a^A \tag{3.4c}$$

$$L^I_{jkat} = \sum_{i \in I} u_{ijkat} + \sum_{l \in L} w_{ljkat} - z_{jkat} \qquad \forall j, k, t, a = 1 \tag{3.5a}$$

$$L^I_{jkat} = L^I_{jk,a-1,t-1} + \sum_{i \in I} u_{ijkat} + \sum_{l \in L} w_{ljkat} - z_{jkat} \qquad \forall j, k, t, 2 \leqslant a < a^A \tag{3.5b}$$

$$L^I_{jkat} = L^I_{jk,a-1,t-1} + L^I_{jka,t-1} + \sum_{i \in I} u_{ijkat} + \sum_{l \in L} w_{ljkat} - z_{jkat} \qquad \forall j, k, t, a = a^A \tag{3.5c}$$

$$\sum_{a \in A} L^I_{jkat} \leqslant b_{jkt} \qquad \forall j, k, t \tag{3.5d}$$

$$\sum_{a \in A} z_{jkat} = d_{jkt} \qquad \forall j, k, t \tag{3.6}$$

$$\sum_{j \in J} \sum_{k \in P} \sum_{a \in A} u_{ijkat} + \sum_{l \in L} \sum_{k \in P} \sum_{a \in A} v_{ilkat} \leqslant S_i \sum_{r \in R(i)} y_{rt} \qquad \forall i, t \tag{3.7}$$

$$\sum_{s \in S} h_{sit} x_s \leqslant \sum_{r \in R(i)} y_{rt} \qquad \forall i, t \tag{3.8}$$

$$y_{rt} \leqslant \sum_{r^p \in R^P(r)} y_{r^p t} \qquad \forall r, t \tag{3.9}$$

$$x_s, y_{rt} \in \{0, 1\} \qquad \forall s, r, t \tag{3.10}$$

$$u_{ijkat}, v_{ilkat}, w_{ljkat}, z_{jkat}, L^F_{ikat}, L^T_{lkat}, L^I_{jkat} \geqslant 0 \qquad \forall i, j, k, a, t \tag{3.11}$$

The objective is to minimize the total cost. The first term represents the costs of schedules, and each team has a given number of sequences. There is a particular cost for each schedule, including harvesting, forwarding, travelling, and costs for moving equipment. The second term expresses costs associated with road opening and maintenance. The third, fourth, and fifth terms in the objective function correspond to transportation costs; from areas to industries, from areas to terminals, and from terminals to industries. The last three terms in the objective function represent storage costs for logs kept at the harvest areas, terminals, and industries.

Constraints (3.1) state that each area can be harvested at most once. Constraints (3.2) specify that each team will be given exactly one of its schedules. The storage balances for logs kept at harvesting areas are described by constraints (3.3a), (3.3b), and (3.3c), where the storage of each assortment and age at the end of each time period are determined. Constraints (3.3a) represent age class 1, which corresponds to products harvested in current time period t. The number of age classes considered is a^A. The highest age class corresponds to timber that is stored a^A time periods or more. After a^A periods the additional reduction of value is 0. The storage of age class a^A is described by constraints (3.3c). Constraints (3.3b) specify the storage level of all age classes except age classes 1 and a^A. Constraints (3.4) are similar to (3.3) but specify the storage levels at terminals. Storage balances at industries are specified by constraints (3.5a), (3.5b), and (3.5c). The storage level at industries is limited through constraints (3.5d). In addition, constraints (3.6) assure that demands at industries will be satisfied.

Constraints (3.7) and (3.8) require at least one road connected to area i to be open if transportation of logs from that area i is done or if area i is harvested, respectively. These two constraints are needed as the logs can be transported from an area in a later time period than the area was harvested. Constraints (3.9) specify the precedence relations between roads. The set $R^P(r)$ is the set of roads, connecting road r to the next level (or higher level, if the state road is the highest) in the road hierarchy. This set is empty for those roads, which are directly connected to the state road. Finally (3.10) and (3.11) define variable bounds.

Generation of harvesting sequences

The number of possible schedules for each team is very large. Areas are spread out all over the district and each team has different home bases. It is reasonable to assume that short travel and moving distances will be more efficient than long travels. We generate a limited number of schedules where those with short distances are more likely to be constructed. The probability of each area to be the first in the schedule is inversely proportional to the travel distance for that combination of area team. The probability of each area to be the next area in the sequence is

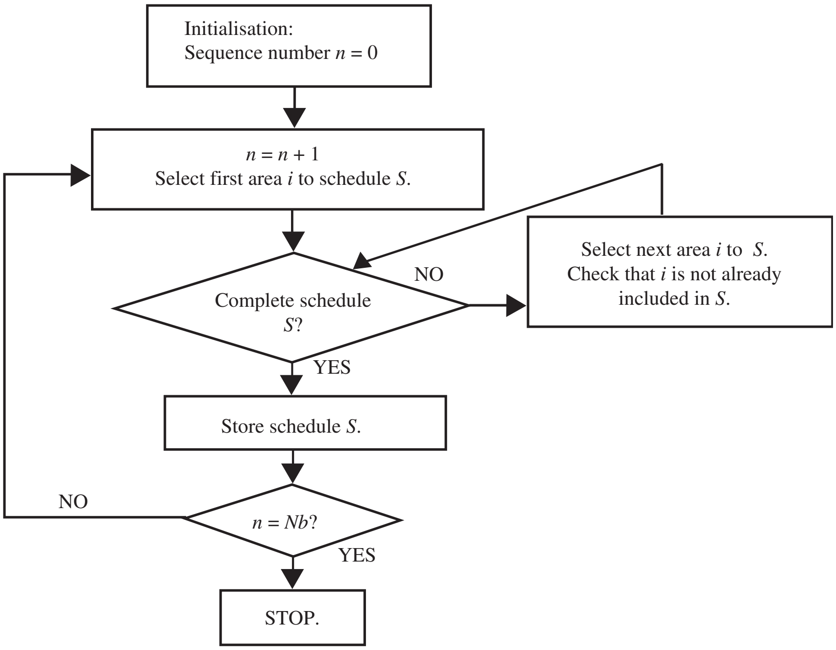

Fig. 3. The flowchart of the algorithm to generate a specified number (Nb) of sequences S for each team.

inversely proportional to the moving distance for that combination of harvest areas. The algorithm that generates a fixed number (Nb) of sequences for each team is described in Fig. 3. During this procedure, the cost of each sequence and the associated data (the corresponding column) are calculated. We have also tried an approach where schedules are randomly generated without any concern to distances, without any improvement. There might be a very small possibility that the same column is generated twice, but this has no effect on the solution approach and is not worth checking.

Case studies

The case study comes from Holmen Skog, a major Swedish company, and its Bergsjö district in the Iggesund region. Holmen is one of the largest forest owners in Sweden with more than one million hectares of forest. Holmen Skog is responsible for providing raw material to the Swedish industries included in Holmen, consisting of a group of companies, where both saw-, pulp- and papermills are included. The Iggesund region is divided into six districts. The annual harvested volume at Bergsjö district is about 250 000 m^3.

We have constructed two basic cases corresponding to the scenarios during two different planning periods in 2000, denoted case A and case B respectively. Cases A^W and B^W correspond to case A and B without usage of age classes (as will be the case during winter). The total planning horizon is five weeks. There are five teams working full-time. The harvesting plan includes harvest crews who work full-time, with the possibility of using an extra resource part-time. The total

Table 3
Information about the size of cases

Aspect	Case A	Case B	Case A^w	Case B^w
Number of areas	34	41	34	41
Number of industries	5	5	5	5
Number of terminals	2	2	2	2
Number of crews	6	6	6	6
Number of assortments	5	5	5	5
Number of age classes	5	5	–	–
Number of time periods	5	5	5	5
Number of Roads	17	21	17	21

supply of different assortments is about 1.5 times the demand. There are five industries. All cost functions and other data needed are obtained from The Forest Research Institute of Sweden. A more detailed description of data and methods for gathering and coordinating these are presented in Bergström, Forsberg and Karlsson (2002). Some basic data from the cases are summarized in Table 3.

Numerical tests

We have used two solution approaches. First we used CPLEX 7.0 and solved the model directly. Second we used a heuristic approach with the purpose to find feasible solutions fast. To evaluate the quality of the LP solutions we have tested a pseudo column generation approach for the linear relaxation of the problem, in order to study the effect of the number of schedules used.

All the computations were performed on a 1.7 GHz Intel workstation, with 1 GB RAM memory. CPLEX was also used for the LP problems obtained in the heuristic solution procedures and in the column generation tests. The mathematical model was implemented using AMPL, see e.g. Fourer, Gay and Kernighan (1993), which is a modeling language system that includes possibilities for developing executable scripts. The harvesting sequences and associated data are generated through a C++ program.

Direct solution

With a limited number of harvesting sequences, the problem can be solved directly using CPLEX. The results are summarized in Table 4.

The solution time differs for the different test examples. The last instance (example 10) takes about six hours to solve. A larger number of schedules generally seems to provide better objective value but requires more computation time. The solution time is within reasonable time limits for these cases.

Heuristic solution approach

We were interested to include more schedules and still find good solutions quickly. From this point of view, it is of interest to develop heuristic solution procedures. Fixing variables in a greedy

Table 4
Results of different test examples applying CPLEX directly (with default settings)

Ex	Case	Sequences	Continuous variables	Constraints	CPU (min)	B&B nodes	Objective value (SEK)
1	A	6000	36,750	5470	6.3	443	1,613,260
2	A	7200	36,750	5470	183	24,667	1,448,400
3	A	8400	36,750	5470	138	14,075	1,446,942
4	A	9000	36,750	5470	200	23,973	1,446,942
5	A	12 000	36,750	5470	120	23,397	1,441,685
6	A^W	12 000	7350	1370	65	10,663	1,390,517
7	B	6000	43,750	6352	13.3	1922	1,791,386
8	B	9000	43,750	6352	233	21,645	1,503,370
9	B	12 000	43,750	6352	148	25,132	1,498,353
10	B^W	12 000	8750	1552	367	72,358	1,434,770

Note: The number of binary variables in the problems is equal to the number of sequences plus the number of binary variables corresponding to road-opening decisions. The number of 0/1 variables associated with roads is 100 in all tests. Ex = Problem number

fashion has been successfully applied to, for example, the crew pairing problem (see Marsten, 1994). The solution approach to the harvest problem in Weintraub et al. (1995) is based on repeated solution of the linear relaxation and fixing integer variables. We have tested a heuristic based on branch-and-bound (Wolsey, 1998) and constraint branching (Ryan and Foster, 1981). We produced a limited branch-and-bound tree and performed only a depth-first search, where the 1-branch (the branch where a variable is fixed to value 1) is chosen. McNaughton (1998) presents an application where constraint branch is applied to harvest planning. Each branching decision deals with a subset of the variables, instead of one single variable as in classical variable branching. Binary variables (in our case, variables representing schedules) are divided into two disjunctive subsets, where the sum of the variables is forced to be equal to 0 or 1 in the two subproblems obtained after branching. In our case, constraint branching strategy gives a more balanced branch and bound tree. Unbalanced tree structure has been proven in many cases to be ineffectively searched in comparison to more balanced trees (Wolsey, 1998). We have used two different approaches to divide schedules into subsets within the branching. In both approaches, branching was based on harvesting decisions concerning one area in a particular time period. We also note that variable branching is a special case of constraint branching where one subset only contains one variable.

The solution procedure involves solving a sequence of LP problems. First, the permanent teams working full time are considered, and secondly teams working part time and road decisions. The heuristic solution procedure is illustrated in Fig. 4.

The first LP solution provides a lower bound of the optimal objective value. As we limited the search to depth search in the 1-branch we are not guaranteed an optimal integer solution, nor even a feasible solution. The reason is that we only explored a part of the entire B&B tree. The method can also be viewed as an integer allocation procedure, in which binary variables are fixed and LP problems are repeatedly resolved. Another approach is to further generate schedules within the heuristic procedure. This can be viewed as applying branch and price (Barnhardt et al., 1998). Here, pricing means that we generate additional columns or schedules in each node of the branch-and-bound tree. It is also possible to apply the approach until all permanent crews have a fixed

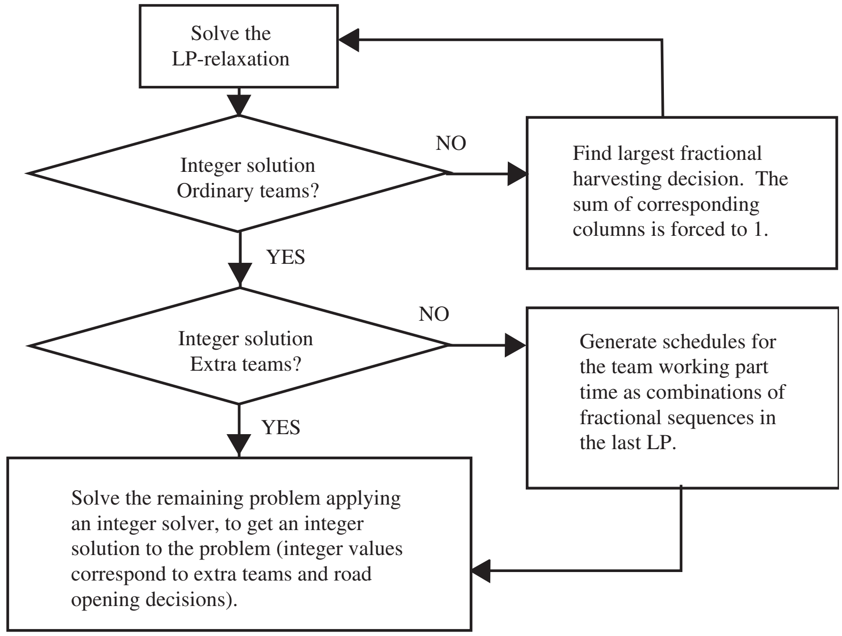

Fig. 4. Flow chart of the heuristic solution procedure.

schedule. The remaining part of the model is much smaller, and we can use a commercial MIP to solve this.

We have implemented two different strategies for implementing constraint branching, both based on a harvesting decision concerning an area i in a time period t. These are denoted Strategy 1 and 2 respectively and are described below.

Strategy 1

In the first strategy, branching is based on the largest fractional *part* of area i that is *harvested* in a time period t, according to the LP solution. For each combination of area and time period (i,t), we estimate the harvesting in proportion to maximum possible amount harvested with respect to possible schedules, h_{it}. Let:

SP = the set of possible schedules s.

SI = the set of schedules where area i is harvested in time period t.

Then,

$$h_{it} = \sum_{s \in SI} \sum_{k \in P} a_{ikts} x_s / \max(\sum_{k \in P} a_{ikts} | s \in SP) \tag{5.1}$$

The largest fractional harvesting, max $(h_{it}|h_{it}<1)$ and corresponding area i' and time period t' is evaluated. In the next LP, area i' is planned to be harvested to the maximum possible in time period t'. Let:

SPF = the set of possible schedules, where: $a_{i'kt's} = \max\ (a_{i'kt's})$

Then in the next LP problem the following constraint is added:

$$\sum_{s \in SPF} x_s = 1 \tag{5.2}$$

The largest possible amount that can be harvested at a certain area, max (a_{ikts}), must be checked after each branching, as the number of possible schedules decreases in each subproblem.

Strategy 2

In the second strategy, the branching is based on the largest fractional value corresponding to the fact that an area i is *started* to be harvested in time period t, according to the LP solution. For each combination of area and time period (i, t), we estimate the sum of the schedules where area i is started to be harvested in time period t, b_{it}. Let:

SB = the set of schedules, where the harvesting of area i is started in time period t.

Then the value of b_{it} is:

$$b_{it} = \sum_{s \in SB} x_s \tag{5.3}$$

The largest fractional start of harvesting, max ($b_{it} | b_{it} < 1$), and corresponding area i' and time period t' is estimated. In the next LP problem, area i' starts to be harvested in time period t'. Let:

SF = the set of schedules where the harvesting of area i' is started in time period t'.

Then in the next LP problem the following constraint is added:

$$\sum_{s \in SF} x_s = 1 \tag{5.4}$$

An advantage of this way of branching compared with the first is that estimation of the largest fractional value to branch on is more effective, as we do not need to estimate max (a_{ikts}) to calculate the fractional values.

The strategies have been applied to a number of test examples based on case A. We have implemented the strategies in AMPL, and results are summarized in Tables 5 and 6. Here, we have example number ('Ex'), the number of sequences included ('Seq.'), CPU-time (in minutes) to find the optimal integer solution ('IP-CPU'), to obtain the heuristic solution ('H-CPU') and to solve the first LP problem ('LP-CPU'). Furthermore, the number of subproblems solved in the heuristic is given ('NB LP'). The objective values of the optimal solution and the heuristic solution are given in SEK. The objective values are compared and the difference is given as a percentage in the last column.

The first branching strategy is too strong. Using depth first search on the test examples 1–5 did not generate feasible solutions. In Table 5 we give the results for examples 2 and 3 if the heuristic procedure is interrupted after a fixed number of subproblems (which provided feasible solutions). One reason there are problems finding feasible solutions is a combination of the strategy and too few schedules. We interrupt after 7 and 10 subproblems and then use CPLEX to solve the remaining problems.

The results of the second heuristic procedure are summarized in Table 6. In these instances, feasible solutions were found. In example 5, we also present the result when the heuristic procedure is stopped after a fixed number of subproblems (10 and 5).

Table 5
Computational result applying heuristic 1, with fixed number of sub-problems

Ex	Seq.	Optimal solution	IP-CPU (min)	Heuristic solution	H-CPU	NB LP	LP-CPU	H/IP (%)
2	7200	1,448,400	183	1,505,971	34	7	7	+4
3	8400	1,446,942	138	1,479,040	44	10	7.3	+2.2

Table 6
Computational result applying heuristic 2

Ex	Seq.	Optimal solution	IP-CPU	Heuristic solution	Heuristic CPU	NB LP	LP-CPU	Diff % H/IP
2	7200	1,448,400	180	1,469,400	41	15	7	+1.5
3	8400	1,446,942	138	1,486,510	39	13	7.3	+2.7
4	9000	1,446,942	200	1,451,260	37	16	8.3	+0.3
5	12 000	1,441,685	120	1,487,839	50	15	6.8	+3.2
5	12 000	1,441,685	120	1,482,097	40	10	6.8	+2.8
5	12 000	1,441,685	120	1,456,577	51	5	6.8	+1.0

In these examples, the heuristic approach yields solutions with an objective value within a few percent of the optimal solutions. Computational time is more than twice as fast as using CPLEX direct. The extended search for example 5 provides a better result, but we cannot guarantee that the remaining integer problem is easy to solve. The second branching strategy is not as strong as the first. A lower bound of the optimal objective value is obtained in the first step of the heuristic. An estimation is that half of the total time needed to get an integer solution is spent on solving the subproblems. Also, AMPL is not efficient when it comes to computational speed. To implement the heuristic in C and CPLEX directly would gain a factor of decrease in CPU time.

Column generation

To increase the number of schedules we apply a pseudo column generation approach. Column generation, developed by Dantzig and Wolfe (1960), is a common method for solving linear problems with a large number of variables (Lasdon, 1970). First a restricted LP problem with a limited number of variables is solved. Then, dual information from the LP solution is used to set up a subproblem to find the best negative reduced cost column. If no column with a negative reduced cost exists, optimality is verified. Column generation is essentially a technique for LP problems. An integer problem can be solved by branch and price (Barnhardt et al., 1998). This involves the classical branch-and-bound method, in which new columns are generated within the branch-and-bound tree.

We use a pseudo column generation approach to make an evaluation about possible improvement of the objective with more schedules. The procedure is based on a large pool of columns generated *a priori*, according to the procedure described in Fig. 3. The pseudo column

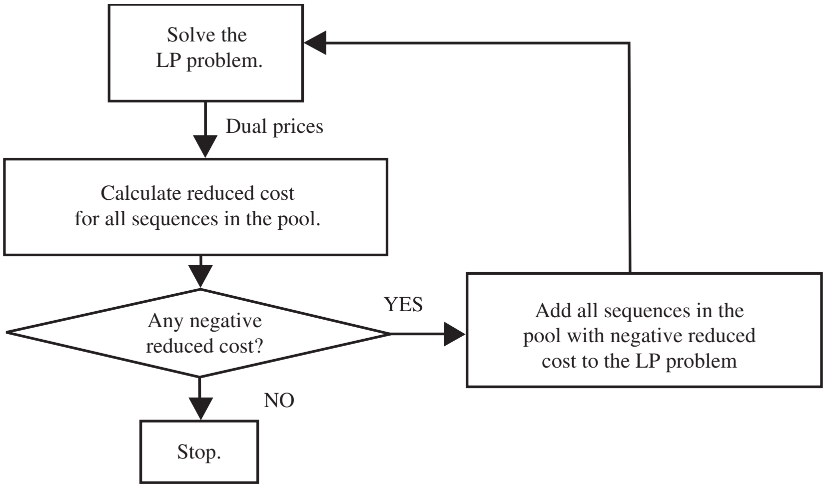

Fig. 5. The pseudo column generation procedure.

Table 7
Results from the column generation test

	Test 1 (Number of sequences in the pool: 150 000)			Test 2* (Number of sequences in each pool: 150 000)		
It.	Sequences in LP	Objective value LP	Nb of seq. Neg. RC	Sequences in LP	Objective value	Nb of seq. Neg. RC
1	3000	1,431,360	3200	3000	1,431,360	3200
2	6200	1,410,046	214	6200	1,410,046	228
3	6414	1,409,467	127	6428	1,405,885	120
4	6541	1,409,463	6	6548	1,404,071	48
5	6547	1,409,463	0	6596	1,403,858	89
6				6685	1,401,643	60
7				6745	1,400,835	31
8				6776	1,400,660	47
9				6823	1,400,581	29
10				6852	1,400,367	20

Notes: It. = iteration number; Sequences in LP number of sequences including variables representing harvest schedules; Nb. of seq. Neg. RC = number of sequences with negative reduced cost in the pool, estimated based on dual variables obtained from the corresponding LP problem.

generation approach is summarized in Fig. 5. First, a large number of feasible schedules are generated. In the first LP problem, a restricted number of these are included. Then the LP problem is solved repeatedly where schedules from the pool with negative reduced costs are added in each step.

The results are summarized in Table 7. In Test 1, the same pool consisting of 150 000 feasible schedules was used throughout the procedure. In each step, all sequences in the pool with a negative reduced cost were added. After six iterations, we did not get any improvements of the objective value. In Test 2, a new pool of 150 000 sequences was considered in each iteration. In each step, all schedules from the active pool, with negative reduced cost were added. It is hard to estimate the total number of columns considered as they are regenerated in each step. The second test was stopped after 10 iterations. The results are summarized in Table 7, where we have the iteration number ('It.'), the number of sequences (variables representing harvest schedules) included in the corresponding LP problem ('Sequences in LP'), the objective value of the LP problem ('Objective value LP') and the number of sequences with negative reduced cost in the pool, estimated based on dual variables obtained from the corresponding LP problem ('Nb. of seq. Neg. RC').

The objective value in Test 1 decreased 1.5% using 150 000 sequences as compared to 3000. In Test 2, the total number of considered columns is larger than 150000 as we generate a new set of 150 000 columns in each iteration. The objective value decreased 2%. A conclusion is that it would be an advantage to include more columns in the suggested IP model, at least to improve the lower bound for the LP relaxation.

Concluding remarks and future work

The developed model and solution process are important parts of a future decision support system for short-term harvesting planning. With such a system, it is possible to quickly generate solutions and test different scenarios. Critical staff can then spend more time on other more qualitative aspects of the forest management. At present, harvesting operations are scheduled manually and planning is strongly dependent on the senior planner at the district. It is very difficult to include all aspects of the decision problem; the age of storage is difficult to consider; manual solutions tend to consider the current problems, e.g. if an industry is lacking some assortment; it is very difficult to include the age of the logs, if so it is merely to remove logs getting too old. This often leads to plans with emergency changes just to satisfy an urgent need.

The proposed model integrates planning of the different steps from harvesting in the forest, to the delivery of logs at the industries. Integrated planning leads to decreased total cost. Furthermore, it is possible to analyze the effect of different changes in the premises, such as changes in demand or production capacities. The model is supposed to be used in a rolling time horizon where a new problem is solved each week. With such an approach, the end states (from one solution) are not critical. This is otherwise a concern with fixed time periods. Solutions were checked and evaluated by persons at the Forest Research Institute of Sweden and planners at Holmen Skog. These evaluations show that the solutions of the model correspond to efficient plans, where the requirements on the plan are satisfied. They do consider an even flow of assortments and the age of logs better than manual plans.

The contributions of this model with respect to earlier work are that this model is detailed and integrates decisions about harvesting, scheduling of harvest crews, transportation, and detailed storage. The model requires information that now is available from customer databases. The resulting mixed integer linear problem is very large. By using a commercial solver it is possible to solve the problem within a practical time limit with a limited number of sequences. We note that it is the optimal solution of a restricted problem as all possible schedules is not included. A developed heuristic provides high-quality solutions within less than half the solution time obtained by the commercial solver. This can be reduced significantly by moving from the general modelling language AMPL to, e.g., a C-implementation.

It is of further interest to include more schedules and to develop the heuristic to guarantee feasible solutions. One approach would be to use branch and price where a subproblem can be solved in order to generate new columns in the branch-and-bound tree. This is likely to be a constrained shortest path problem similar to those obtained in, e.g., airline crew scheduling.

References

Barnhart, C., Johnson, E.L., Nemhauser, G.L., Savelsbergh, M.W.P., Vance, P.H., 1998. Branch-and-price: Column generation for solving huge integer programs. *Operations Research* 46, 316–332.

Bergström, J., Forsberg, D., Karlsson, J., 2002. Turordningsplanering. *The Forest Research Institute of Sweden, Working paper*, (in Swedish).

Brunberg, T., 2001. Flexibel drivning på Östgården hos Norrskog. *The Forest Research Institute of Sweden, Working paper 472-2001*, (in Swedish).

Carlsson, D., Rönnqvist, M., 1998. Wood flow problems in the Swedish forestry. *Department of Mathematics, Linköping Institute of Technology, LiTH-MAT-R-1998-16*, Sweden.

Epstein, R., Morales, R., Seron, J., Weintraub, A., 1999a. Use of OR Systems in the Chilean Forestry Industry. *Interfaces* 29, (1), 7–29.

Epstein, R., Weintraub, A., Chevalier, P., Gabarro, J., 1999b. A system for short term harvesting. *European Journal of Operational Research* 119, 427–439.

Fourer, R., Gay, D.M., Kernighan, B.W., 1993. *AMPL – A Modeling Language For Mathematical Programming*. Scientific Press.

Jarmer, C., Sessions, J., 1992. Logger-PC for improved logging planning. *Proceedings of Planning and Implementing Future Forest Operations*, International Mountain Logging and 8th Pacific Northwest Skyline Symposium, 14–16 December 1992, Bellevue, Washington, College of Forest Resources, Box 352100, University of Washington, Seattle, Washington, pp. 241–247.

Karlsson, J., Rönnqvist, M., Bergström, J., 2002. Annual harvest planning. *Department of Mathematics, Linköping Institute of Technology, LiTH-MAT-R-2002-15*, Sweden.

Lasdon, S.L., 1970. *Optimizing theory for large system*. Macmillan Publishing Co., New York.

Marsten, R., 1994. Crew planning at Delta Airlines. Presentation at Mathematical Programming Symposium XV, Ann Arbor, MI.

McNaughton, A., 1998. Long-term scheduling of harvesting with adjacency and trigger constraints. Ph.D. Thesis, University of Auckland, Auckland, New Zealand.

McNaughton, A., Rönnqvist, R., Ryan, D., 2000. A model which integrates strategic and tactical aspects of forest harvesting. *Proceedings of System Modelling and Optimization Methods, Theory and Applications*. 19th IFIP TC7 Conference on System Modelling and Optimization, 12–16 July 1999, Cambridge, UK.

Nelson, J., Brodie, J.D., 1990. Comparison of random search algorithm and mixed integer programming for solving area-based forest plans. *Canadian Journal of Forestry Research* 20, 934–942.

Ryan, D.M., Foster, B.W., 1981. An Integer programming approach to scheduling. *Computer Scheduling of Public Transport*, North Holland, Amsterdam, pp. 269–280.

Weintraub, A., Cholaky, A., 1991. A hierachical approach to Forest Planning. *Forest Science* 37, (2), 439–460.

Weintraub, A., Jones, G., Meacham, M., Magendzo, A., Magendzo, A., Malchuk, D., 1995. Heuristic procedures for solving mixed-integer harvest scheduling-transportation planning models. *Canadian Journal of Forest Research* 25, 19, 1618–1626.

Wolsey, L.A., 1998. *Integer Programming*. Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, Chichester.

Yoshimoto, A., Brodie, J.D., Session, J., 1994. A new heuristic to solve spatially constrained long-term harvest scheduling problems. *Forest Science* 40, 365–396.