Comparing Alternative Formulations for the ARM

Juan Pablo Vielma — Georgia Institute of Technology
Marcos Goycoolea — Universidad Adolfo Ibañez
Alan T. Murray — Ohio State University
Andres Weintraub — Universidad de Chile

Abstract

In this article we study the effectiveness of alternative integer programming formulations for area constrained harvest scheduling, known as the area restriction model (ARM). Empirical assessment of the effect of area constraints on the optimal objective value of these alternative approaches is carried out, focusing on the root Linear Programming relaxation. We also examine how this relates to the effectiveness of these formulations when solved by a commercial solver, CPLEX.

Introduction

In this article we discuss harvesting models in the context of the area restriction model (ARM) (Murray 1999). The ARM includes aggregation of small basic cells into bigger clear-cut sections and hence is more difficult to solve than the URM based models. However, the ARM formulation leads to superior solutions (Murray and Weintraub 2002), making it more appealing for harvest scheduling, provided that supporting data exists. The ARM can consider area constraints that extend multiple periods and the length of this extension is usually denoted as the green-up length or period.

Given the difficulties in solving ARM formulations, most solution approaches until recently have relied on heuristics, like Tabu search or Simulated Annealing (Lockwood and Moore 1993, Barrett et al 1998, Richards and Gunn 2000, Boston and Bettinger 2002, Caro et al. 2003). In recent years, however, different exact approaches have been proposed as integer programs. One approach, called the Cell Approach, is based on enumerating all harvesting blocks that violate the area constraints (McDill et al 2002, Crowe et al 2003, Gunn and Richard 2005). A second approach, called the Cluster Approach, is based on enumerating all harvesting blocks that do not violate area constraints, and preventing simultaneous harvest of any two adjacent or intersecting blocks (Martins et al 1999, McDill et al 2002, Goycoolea et al 2005, Vielma et al 2005).

The advantages and disadvantages of these formulations are discussed in Goycoolea et al. (2007). One of the issues discussed in Goycoolea et al. (2007) is that theoretically the linear programming (LP) relaxation of the cluster approach is tighter than the cell approach. This is confirmed empirically for single period instances and multi-period instances without side constraints. It is also shown that the difference between the LP relaxations of the two formulations can be quite large. However, when side constraints, such as volume flow or average ending age, are added this difference becomes smaller.

When green-up periods greater than one are considered, however, the difference becomes large again. Furthermore, the cluster approach presents a large computational advantage over the cell approach when its LP relaxation is much tighter. For this reason it is important to understand why the difference between the LP relaxations is large in some instances and small in others. The results from Goycoolea et. al. (2007) suggest that the difference between the LP relaxation is larger when area constraints are the main restrictive factor and that it decreases when other side constraints are included. In this paper we further explore this phenomenon by empirically studying the effect of area constraints on the optimal objective value of different variations of the ARM.

The remainder of the paper is organized as follows. In sections 2 and 3 descriptions of the ARM approach are given. Side constraints used here and green-up constraints are detailed in sections 4 and 5. Section 6 presents computational results. Finally, section 7 gives comments and conclusions.

The ARM

Typically forest models are made up of basic cells of 1 to 20 hectares, which constitute the minimal units over which decisions are made. Given that typical maximum opening size of clearings usually range from 40 to 120 hectares, these cells can be merged for harvesting purposes into blocks. The ARM then is structured to make decisions about when to harvest each cell, if at all, in such a way that contiguous clear cut areas do not exceed a prescribed maximum area.

Figure 1 shows an example of a small forest composed of 11 cells, where cells 1 and 4 form block 1, cells 7, 8 and 11 form block 2 and cell 6 forms block 3. Since these blocks are not adjacent, and we assume them of size below the maximum opening size, harvesting these blocks in period one constitutes a feasible solution with respect to the maximum area constraint.

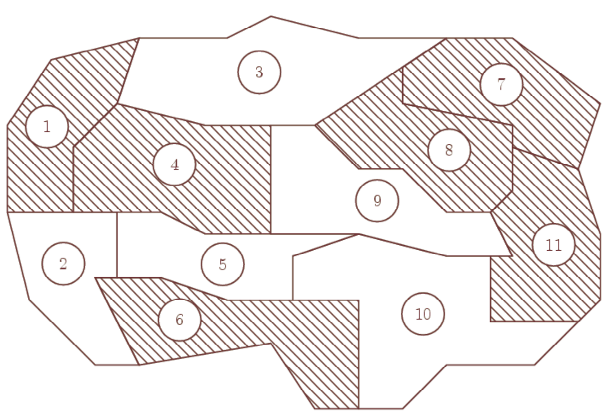

Figure 1

There can also be additional constraints that ensure a smooth timber production through time as well as ending age conditions at the horizon.

Modeling Approaches

We can represent the relationship of cells in a graph G=(V,E), where cells are represented by vertices V and vertices are joined by an edge E if the two corresponding cells are adjacent.

Cell Approach

For the cell approach, binary variables $y_{v,t}$ represent harvesting cell v in time period t (McDill et al 2002, Crowe et al 2003, Gunn and Richards 2005). Area constraints are enforced by using the concept of minimally infeasible clusters. Minimally infeasible clusters are sets of contiguous cells whose total area exceeds the maximum clear-cut area, but that become discontiguous or comply with the maximum clear-cut area if any cell is removed from the set.

Denoting the set of minimal infeasible clusters as Λ, the cell formulation is:

$$\min \sum_{t \in T} \sum_{v \in V} p_{v,t} y_{v,t} \qquad (1)$$

Subject to

$$\sum_{t \in T} y_{v,t} \leq 1 \qquad \forall v \in V \qquad (2)$$

$$\sum_{v \in D} y_{v,t} \leq |D| - 1 \qquad \forall D \in \Lambda, \forall t \in T \qquad (3)$$

$$y_{v,t} \in \{0,1\} \qquad \forall v \in V, \forall t \in T \qquad (4)$$

where $p_{v,t}$ is the revenue obtained if cell v is harvested in period t.

Cluster Approach

An alternative ARM approach is to focus on feasible clusters. A feasible cluster is a group of contiguous cells whose combined area is less than or equal to the maximum allowed clear-cut area. In the cluster approach, variables $x_{C,t}$ represent harvesting feasible cluster C in time t (see Martins et al. 1999, Goycoolea et al. 2005, Vielma et al. 2007). For a green-up of 1 period, area constraints are enforced by forbidding adjacent or intersecting clusters from being harvested in the same period, and can be easily extended to green-ups greater than one.

If we let Ω be the set of all feasible clusters, the cluster formulation is:

$$\max \sum_{C\in\Omega}\sum_{t\in T} p_{C,t}x_{C,t} \qquad (5)$$

Subject to

$$\sum_{t\in T}\sum_{C\in\Omega(v)} x_{C,t} \le 1 \qquad \forall v\in V \qquad (6)$$

$$\sum_{C\in\Omega(K)} x_{C,t} \le 1 \qquad \forall K\in\Pi, \forall t\in T \qquad (7)$$

$$x_{C,t}\in\{0,1\} \qquad \forall C\in\Omega, \forall t\in T \qquad (8)$$

where $p_{C,t}$ is the revenue obtained if cell v is harvested in period t, Π is the set of all maximal cliques in G and $\Omega(K)$ is the set of all clusters that intersect clique K.

Side Constraints

One side constraint that is usually added to the ARM is to require a non-declining flow of volume of timber. These constraints can be implemented in the cell aggregation formulation by adding:

$$\begin{aligned} \sum_{v\in V}\alpha_{v,t+1}y_{v,t+1} &\le U_t\sum_{v\in V}\alpha_{v,t}y_{v,t} \qquad \forall t\in 1,\ldots,T-1 \\ \sum_{v\in V}\alpha_{v,t+1}y_{v,t+1} &\ge L_t\sum_{v\in V}\alpha_{v,t}y_{v,t} \qquad \forall t\in 1,\ldots,T-1 \end{aligned} \qquad (9)$$

where $\alpha_{v,t}$ denotes the volume of timber obtained when harvesting cell v at time t.

This can be done for the cluster packing formulation by adding:

$$\begin{aligned} \sum_{C\in\Omega}\alpha_{C,t+1}x_{C,t+1} &\le U_t\sum_{C\in\Omega}\alpha_{C,t}x_{C,t} \qquad \forall t\in 1,\ldots,T-1 \\ \sum_{C\in\Omega}\alpha_{C,t+1}x_{C,t+1} &\ge L_t\sum_{C\in\Omega}\alpha_{C,t}x_{C,t} \qquad \forall t\in 1,\ldots,T-1 \end{aligned} \qquad (10)$$

where $\alpha_{C,t}=\sum_{v\in C}\alpha_{v,t}$ equals the volume of timber obtained when harvesting cluster $C\in\Omega$ at time t.

Another side constraint is to require that the forest have at the end of the planning horizon an average age of at least $\overline{G}$ years. These constraints can be modeled in the cell approach by adding:

$$\sum_{t}\sum_{v\in V}(g_{v,t}-\overline{G})y_{v,t}\le 0 \qquad (11)$$

and in the cluster model by adding,

$$\sum_{t}\sum_{C\in\Omega}(g_{C,t}-|C|\overline{G})x_{C,t}\le 0 \qquad (12)$$

where $g_{i,t}$ is the age of basic cell i∈V at the end of the planning horizon if it is harvested in time t and $g_{C,t}=\sum_{i\in C}g_{i,t}$ for each cluster C∈Ω and time t

Green-up Constraints

Green-up constraints extend area constraints over multiple time periods, denoted Δ periods. This Δ is known as the green-up periods.

Green-up constraints are enforced by introducing additional binary variables $z_{v,t}$ for each v in V and each t in {1,…,T} and modifying constraints. The idea is that variables $z_{v,t}$ will now indicate if cell v is harvested in period t and variables $y_{v,t}$ and $x_{C,t}$ will indicate if cell v and cluster C are considered clearcut areas in period t respectively. The modification of the constraints is as follows. For the cell approach constraints (2) need to be replaced by:

$$y_{v,t}=\sum_{q=t-\Delta+1}^{t}z_{v,q} \qquad \forall t\in T,\forall v\in V \qquad (13)$$

For the cluster approach constraints (6) need to be replaced by:

$$\sum_{\{C:v\in C\}}x_{C,t}=\sum_{q=t-\Delta+1}^{t}z_{v,q} \qquad \forall t\in T,\forall v\in V \qquad (14)$$

For both approaches it is necessary to add constraints:

$$\sum_{t=1}^{T} z_{v,t} \leq 1 \qquad \forall v \in V \qquad (15)$$

In addition, changes need to be made to both models, since the objective function, volume flows, and average ending age constraints should be stated in terms of the z variables as opposed to the y and x variables. For example, objective functions (1) and (5) should be replaced by:

$$\max \sum_{t\in T} \sum_{v\in V} p_{v,t} z_{v,t} \qquad (16)$$

Computational Results

In this section we study the LP gaps and solve times of the cluster and cell formulations and relate them to the effect of maximum clear-cut area constraints on the optimal objective value.

The analysis uses data from the El Dorado National Forest in northern California, which is available at the *FMOS (2006) web site. This forest consists of 1,363 cells ranging from 10 to 116.35 acres, for a total area of 52,255.45 acres. We used the area, age and timber volume information for each cell and assumed revenue was proportional to the timber volume. We additionally considered a discount rate of 3% per period, and two cells were defined as adjacent if they touched. As side constraints, we included volume flow constraints with $L_t = 1 - 15/100$, and $U_t = 1 + 15/100$, for all $t \in T$ and average ending age constraints of 40 years.

All runs were made on a Pentium IV (Xeon) running at 2.40 GHz and with 2 GB of RAM running Linux. All programs were written in the C++ programming language, and ILOG CPLEX 9.0 was utilized for all linear and integer programs solves.

Root LP Gaps and Solve Times

For each planning horizon we considered up to four instances; one without using any side constraints (labeled as “T= ?, no side const”, where ? is the length of the planning horizon), one using both volume flow and average ending age constraints (“T= ?”), and two using volume flow, average ending age, and green-up constraints (“T=?, green-up=?”, where the last parameter indicates the length of the green-up period). Each problem was solved with a time limit of 10,000 seconds and for each problem we present the root LP gap, the best gap obtained in the allotted time and the solve time in parenthesis if optimality (considered as a 0.01% gap) was reached. The root LP gap is the ratio between the value of the LP relaxation and the value of the best known feasible solution for the corresponding problem. If x is the value of the best known solution for a given instance, and r is the value of the LP relaxation obtained using a given formulation, then the corresponding root LP gap would have value (r/x-1)*100. For example, the best known feasible solution for El Dorado (T=3) has value x=2612790 and the LP relaxation for the cluster formulation has value r = 2617280. Thus, the value of root LP gap is (2617280/2612790-1)*100 = 0.17. Note that for some instances the best known solution was not found by either of the formulations in the allotted time. Given an incomplete run of a branch and bound algorithm, let zu be the value of the current upper bound and zl be the value of the best known feasible solution. We compute gap as (zu/zl-1)*100. If no feasible solution was found, we indicate this with a dash in the corresponding column The results are illustrated in Table 1.

		Cell		Cluster	
		Root lp gap (%)	Final gap (Opt time)	Root lp gap (%)	Final gap (Opt time)
El Dorado	T=3, no side const	5.14%	2.99%	0.43%	(659 s)
	T=3	2.75%	1.16%	0.17%	(1269 s)
	T=3, greenup=2	9.02%	7.33%	0.34%	(5154 s)
	T=5, no side const	3.12%	1.51%	0.28%	(903 s)
	T=5	0.27%	0.18%	0.09%	0.15%
	T=5, greenup=2	5.30%	6.20%	0.64%	6.65%
	T=5, greenup=3	9.07%	11.33%	0.54%	0.24%
	T=12, no side const	0.70%	0.24%	0.00%	(186 s)
	T=12	0.5%	1.15%	0.5%	3.08%
	T=12, greenup=2	0.88%	2.67%	0.74%	6.37%
	T=12, greenup=3	2.56%	-	1.39%	14.81%

Table 1. Root LP gap’s and solve times.

We can see that the LP relaxation of the cluster approach is always better than the LP relaxation of the cell approach. The difference is much bigger though for the instances without side constraints or with green-ups close to the total number of periods. We can also see that the cluster approach's tight LP relaxation usually yields a much better performance than the cell approach. Still, in the instances in which the root gap are practically the same (such as T=12 and T=12, greenup=2) the cell approach can yield a better performance than the cluster approach. It seems then that the cluster approach's performance advantage is tied to tight root LP gaps and that this advantage is lost for some instances in which the cell approach also has a tight root LP gaps. In the following section we study an aspect that could explain why the cell approach can have root LP gaps practically as small as the cluster approach for some instances.

Effect of Area Constraints

In this section we study the effect of maximum area constraints over the optimal objective value. To achieve this we set out to measure the increment in the optimal objective function value when maximum area constraints are removed. We chose the cell approach for these experiments as removing the area constraints for this formulation is achieved easily by removing constraints (3).

For each instance we measure the increment in the optimal objective value when maximum area constraints are removed as follows. Let w be the optimal objective value of the corresponding cell approach formulation and let s be the optimal objective value of the problem obtained by removing constraints (3) from the corresponding cell approach formulation. We measure the increment as

$$\text{Increment}=(s/w-1)*100.$$

The results are illustrated in Table 2. We note that because we could not solve all the problems to optimality there is a small error in the data of Table 1. The positive additive error for most of the reported increments is smaller than 0.1% tough. The exceptions are marked with one or two stars (*). One star (*) means the error is larger than 0.1% but smaller than 1% and two stars (**) means the error is larger than 1% but lower than 3%. For example for "T=3, no side constraints" the increment is between 24.4% and 24.5%, for "T=12" the increment is between 0% and 1% and for "T=12, green-up=2" the increment is between 0.7% and 3.7%.

		Increment
El Dorado	T=3, no side const	24.4%
	T=3	7.70%
	T=3, greenup=2	38.7%
	T=5, no side const	13.6%
	T=5	1.2%
	T=5, greenup=2	16%*
	T=5, greenup=3	35.4%
	T=12, no side const	3.5%
	T=12	0%*
	T=12, greenup=2	0.7%**
	T=12, greenup=3	4.7%**

Table 2. Increment in the optimal value when area constraints are removed.

From table 2 we see that for 3 and 5 periods without side constraints there is a significant increment of the optimal objective function value when the area constraints are removed. This is not surprising as, in this case the only constraints besides the area constraints are the ones that forbid harvesting a cell more than once (2). What is somewhat surprising is that even without side constraints removing the area constraints for the 12 period instance only increments the optimal objective value by a small percentage. When side constraints are added we see that the increment in the optimal objective function obtained when area constraints are removed is never greater than 8%. Furthermore, for the 12 period instance with side constraints the increment is practically nonexistent. Finally, we see that for green-up>1 the increments are again significant, specially when the green-up is close to the total number of periods. The increment is still very small though for a green-up of 2 periods and a total of 12 periods.

We also see that the instances with very small increment (<2%) coincide with the instances for which the cell approach has a very good root LP gap (<=0.5%). On the other hand, the cluster approach has excellent root LP gaps even when the increment is large. This suggests that the cell approach's main deficiency is that it does not approximate well the effect of area constraints, but that this deficiency is inconsequential when the effect of area constraints is negligible.

Comments and Conclusions

We have empirically studied the effect of area constraints over the optimal objective function value of ARM formulations. We have seen that area constraints can significantly decrease the optimal objective function value for problems without side constraints and problems with side constraints and green-up periods that are not too small compared to the planning horizon. On the other hand, when side constraints are included and the green-up period is small compared to the planning horizon, area constraints have little impact in the optimal objective function value. This illustrates how for some ARM instances area constraints might not be that important and what we denoted as side constraints can be actually be the important constraints.

We have also seen that the instances where area constraints have little impact on the optimal objective, and coincide with the instances for which the cell approach formulation has a root LP gap almost as good as the cluster formulation. These are also the problems where the cell approach's performance is comparable or better than that of the cluster approach. For the instances in which the area constraints have a high impact, the root LP gap and the performance for the cluster approach is much better than that of the cell approach.

The fact that side constraints can be the difficult constraints for the cluster approach has been previously discussed in Vielma et al. (2007). Here it was noted that in 12 period instances, volume constraints could have a big impact on solve times and it was shown that replacing hard volume flow constraints by elastic versions of the constraints could significantly improve them. On the other hand, the cell approach seems to be less sensitive to hard side constraints. Although it did not perform significantly better than the cluster approach for the 12 period instance with side constraints its performance barely deteriorated from when side constraints where not present. This suggests a further study of the sensitivity of the cell approach to side constraints.

A somewhat surprising result is the small impact on the objective value of the constraints for the 15 period instance without side constraints. It seems that for this instance, even in the absence of side constraints, harvesting a cell when it is most valuable can be easily accomplished without violating area constraints. It would be interesting to study what forest characteristics lead to this phenomenon.

References

Barrett, T., Gilless J. and L. Davis. 1998. Economic and Fragmentation Effects of Clearcut Restrictions. *Forest Science* Vol. 44, N°4, 569-577.

Boston, K. and P. Bettinger. 2002. Combining tabu search and genetic algorithms heuristic techniques to solve spatial harvest scheduling problems. *Forest Science*. 48:35-46.

Caro, F., M. Constantino, I. Martins, A. Weintraub (2003). "A 2-Opt Tabu Search Procedure for the Multi-Period Forest Harvesting Problem with Adjacency, Green-up, Old Growth and Even Flow Constraints". *Forest Science*. 49(5) 738-751.

Crowe, K., J. Nelson, and M. Boyland. 2003. "Solving the area-restricted harvest scheduling model using the branch and bound algorithm". Can. J. For. Res./Rev. can. rech. for. 33(9): 1804-1814.

Forest Management Optimization Site (FMOS). 2006. Richards, E.W. Accessible through http://www.unbf.ca/fmos/.

Goycoolea M., Murray A., Barahona F., Epstein R., Weintraub A., 2005, Harvest scheduling subject to maximum area restrictions: exploring exact approaches. Operations Research 53, 490-500.

Goycoolea M., Murray A., Vielma J.P., Weintraub A., 2007, Evaluating Approaches for Solving the Area Restriction Model in Harvest Scheduling, Working Paper.

Gunn, E. A. and E.W. Richards. 2005. "Solving the adjacency problem with stand-centred constraints". Canadian Journal of Forest Research. 35(4) 832-842

Lockwood, C. and T. Moore. 1993. Harvest scheduling with spatial constraints: a simulated annealing approach. *Can. J. For. Res.* 23: 468-478.

Martins I, M. Constantino and J. Borges. 1999. "Forest management models with spatial structure constraints." Working paper 2/99. Centro de Investigaçao Operacional. Universidade de Lisboa. Lisbon, Portugal.

McDill, M. E., S. Rebain, and J. Braze. 2002. "Harvest Scheduling with Area-Based Adjacency Constraints." *Forest Science* 48: 631–642.

Murray, A.T. 1999. Spatial Restrictions in Harvest Scheduling. *Forest Science* 45: 1-8.

Murray A. T. and Weintraub, A. 2002. Scale and unit specification influences in harvest scheduling with maximum area restrictions. *Forest Science* 48:779-789.

Richards, E.W. and Gunn, E.A. 2000. "A model and tabu search method to optimize stand harvest and road construction schedules." *Forest Science* 46: 188-203.

Vielma, J.P., A.T. Murray, M. Ryan, and A. Weintraub. 2007. “Improving Computational Capabilities for Addressing Volume Constraints in Forest Harvest Scheduling Problems”. European Journal of Operational Research 176:1246–1264