CIRRELT

Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

A Stochastic Programming Approach for Production Planning with Uncertainty in the Quality of Raw Materials: A Case in Sawmills

Masoumeh Kazemi Zanjani
Mustapha Nourelfath
Daoud Aït-Kadi

February 2009

CIRRELT-2009-08

Bureaux de Montréal :
Université de Montréal
C.P. 6128, succ. Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince, local 2642
Québec (Québec)
Canada G1K 7P4
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Université Laval · UQÀM · HEC Montréal · École Polytechnique Montréal · Université de Montréal

A Stochastic Programming Approach for Production Planning with Uncertainty in the Quality of Raw Materials: A Case in Sawmills

Masoumeh Kazemi Zanjani[1,*], Mustapha Nourelfath[1], Daoud Aït-Kadi[1]

[1.] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Département de génie mécanique, Pavillon Adrien-Pouliot, Université Laval, Québec, Canada G1K 7P4

Abstract. Motivated by sawmill production planning, this paper investigates multi-period, multi-product (MPMP) production planning in a manufacturing environment with non-homogeneous raw materials, and consequently random processes yields. A two-stage stochastic linear program with recourse is proposed to address the problem. The random yields are modeled as scenarios with stationary probability distributions during the planning horizon. The solution methodology is based on the sample average approximation (SAA) scheme. The stochastic sawmill production planning model is validated through Monte Carlo simulation. The computational results for a real medium capacity sawmill highlight the significance of using the stochastic model as a viable tool for production planning instead of the mean-value deterministic model, which is a traditional production planning tool in many sawmill.

Keywords. Production planning, random yield, sawmill, stochastic programming, sample average approximation, Monte Carlo simulation.

Acknowledgements. The authors would like to acknowledge the financial support provided by the Forest E-business Research Consortium (FOR@C) of Université Laval, and would like to thank specially Jonathan Gaudreault, Philippe Marier, Sébastien Lemieux, and Christian Rouleau, for their technical support.

---

* Corresponding author: Masoumeh.KazemiZanjani@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec,
Bibliothèque et Archives Canada, 2009

1. Introduction

Most of the production environments are characterized by multiple types of uncertainties. When planned production quantities are released, the outputs are often variable. These uncertainties affect and complicate the production planning and control.

The goal of this work is to address a multi-period, multi-product (MPMP) sawmill production planning problem. In sawmills, raw materials (logs) are classified based on some attributes namely: diameter class, species, length, taper, etc. Logs are broken down into different pieces of lumbers (e.g. 2(in)× 4(in)×8(ft), 2(in)×4(in)×10(ft),…) by means of different cutting patterns. However, due to the non-homogeneity in characteristic of logs, the quantities of lumbers sawn by different cutting patterns (processes yields) are random variables. In fact, due to natural variable conditions that occur during the growth period of trees, it is impossible to anticipate the exact yields of a log. Moreover, as it is not possible in many sawmills to scan the logs before planning, the exact yields of cutting patterns for different log classes cannot be determined in priori. Product demand is considered as a deterministic parameter which is determined based on the received orders. Production planning in sawmills is to determine the optimal quantity of log consumption from different classes and the selection of corresponding cutting patterns to fit against products demands. The part of the demand that cannot be fulfilled on time due to machine capacities, log inventory, and random yield will be postponed to the following periods by considering a backorder cost. The objective is to minimize log consumption cost, as well as products inventory/backorder costs. We are studying a very customer orientated manufacturing environment that wishes to fulfill the demand as much as possible. Regarding to the potential significance of yield uncertainty on the production plan and consequently on the realized total backorder size, obtaining the plans with minimum total backorder size is an important goal of production planning in sawmills.

This production planning problem can be considered as the combination of several classical production planning problems in the literature which have been modeled by linear programming (LP). Product mix problem and a special case of process selection problem (Johnson and Montgomery, 1974; Sipper and Bulfin, 1997) are the two main building blocks of this problem. In Gaudreault et al. (2004), a deterministic LP model was proposed for sawmill production planning by considering the expected values of random processes yields. The production plan proposed by the deterministic model results usually extra inventory of products with lower quality and price while backorder of products with higher quality and price. Another approach for sawmill production planning is focused on combined optimization type solutions linked to real-time simulation sub systems (Mendoza et al., 1991; Maness and Adams, 1991; Maness and Norton, 2002). In this approach, the stochastic characteristics of logs are taken into account by assuming that all the input logs are scanned through an X-ray scanner, before planning. Maness and Norton (2002) developed an integrated multi-period production planning model which is the combination of an LP model and a log sawing optimizer (simulator). The LP model acts as a coordinating model that allocates limited resources. A series of dynamic programming sub-problems, titled in the literature as “log sawing optimization models” are used to generate activities (columns) for the coordinating LP based on the products’ shadow prices. Although the stochastic characteristics of logs are considered in this approach, it includes the following limitations to be implemented in many sawmills: logs, needed for the next planning horizon, are not always available in sawmills to be scanned before planning. Furthermore, to implement this method, the logs should be processed in production line in the same order they have been simulated, which is not an easy practice. Finally, scanning logs before planning is a time consuming process in the high capacity sawmills, which delays the planning process.

It has been shown in the literature (see for example Kall and Wallace, 1994; Birge and Louveaux, 1997; Kall and Mayer, 2005) that in mathematical programming models which include random parameters in their right-hand–side and/or technological coefficients, stochastic programming approach results higher quality solutions compared with the mean-value deterministic model. Most of the works in the literature on uncertain production planning are focused on considering random products demands. In Escudero et al. (1993), a multi-stage stochastic programming approach was proposed to address a MPMP production planning model with random demand. Sox and Muckstadt (1996) provided a formulation and solution algorithm for the finite-horizon capacitated production planning problem with random demand for multiple products. Using Lagrangian relaxation, they developed a sub gradient optimization algorithm to solve the problem. In Bakir and Byrune (1998), demand uncertainty in a MPMP production planning model was studied. They developed a demand stochastic LP model based on the two-stage deterministic equivalent problem. In Kazemi et al. (2007), stochastic programming was proposed as one of possible methodologies to address sawmill production planning, while considering random characteristics of logs.

In this paper, a two-stage stochastic program with recourse (Kall and Wallace, 1994; Birge and Louveaux, 1997; Kall and Mayer, 2005) is proposed for sawmill production planning, while considering random characteristics of raw materials and consequently random processes yields. We also propose an approach to model the random processes yields in sawmills. Due to the astronomic number of scenarios for random yields in the two-stage stochastic model, a Monte Carlo sampling strategy, the sample average approximation (SAA) scheme (cf. Shapiro and Hommem-de-Mello, 1998, 2000; Mak et al., 1999) is implemented to solve the stochastic model. The confidence intervals on the optimality gap for the candidate solutions are constructed based on common random number (CRN) streams (Mak et al., 1999). Through Monte Carlo simulation, we compare the stochastic and deterministic sawmill production planning models in terms of the realized backorder size and precision of proposed models. Our computational results involving one medium capacity sawmill, with different demand levels, indicate that the proposed stochastic programming approach serves as a viable tool for sawmill planning by considering random characteristics of logs.

The remainder of this paper is organized as follows. In the next section, we propose a two-stage stochastic linear program for sawmill production planning under uncertainty of processes yields. In section 3, the proposed approach for modeling random processes yields in sawmills is provided. In section 4, we provide the solution methodology for the two-stage stochastic model. In section 5, the proposed validation approach to compare the stochastic and deterministic sawmill production planning models is presented. In section 6, the implementation results of the stochastic model and solution strategy for a realistic scale sawmill are presented. The results of comparison between the plans of stochastic and mean-value deterministic LP models are also reported in this section. Our concluding remarks are given in section 7.

2. Problem formulation by mathematical programming

In this section we first describe the deterministic linear program (LP) formulation for sawmill production planning. Then we develop the proposed stochastic model to address the problem by considering the uncertainty of processes yields.

2.1. The deterministic LP model for sawmill production planning

Consider a sawmill with a set of products (lumbers) P, a set of classes of raw materials (logs) C, a set of production processes A, a set of resources (machines) R, and a planning horizon consisting of T periods. For modeling simplicity, we define a process as a combination of a log class and a cutting pattern. To state the deterministic linear programming model for this production planning problem, the following notations are used:

2.1.1. Notations

Indices

p product (lumber)

t period

c raw material (log) class

a production process

r resource (machine)

Parameters

h_{pt} Inventory holding cost per unit of product p in period t

b_{pt} Backorder cost per unit of product p in period t

m_{ct} Raw material cost per unit of class c in period t

I_{c0} The inventory of raw material class c at the beginning of planning horizon

I_{p0} The inventory of product p at the beginning of planning horizon

s_{ct} The quantity of material of class c supplied at the beginning of period t

d_{pt} Demand of product p by the end of period t

ϕ_{ac} The units of class c raw material consumed by process a (consumption factor)

ρ_{ap} The units of product p produced by process a (yield of process a)

δ_{ar} The capacity consumption of resource r by process a

M_{rt} The capacity of resource r in period t

Decision variables

X_{at} The number of times each process a should be run in period t

I_{ct} Inventory size of raw material of class c by the end of period t

I_{pt} Inventory size of product p by the end of period t

B_{pt} Backorder size of product p by the end of period t

2.1.2. The LP model

$$\text{Minimize } Z = \sum_{p\in \mathrm{P}} \sum_{t=1}^{T} (h_{pt} I_{pt} + b_{pt} B_{pt}) + \sum_{c\in \mathrm{C}} \sum_{t=1}^{T} \sum_{a\in A} m_{ct} \phi_{ac} X_{at} \tag{1}$$

Subject to

Material inventory constraint

$$I_{ct} = I_{ct-1} + s_{ct} - \sum_{a\in A} \phi_{ac} X_{at}, \qquad t = 1,\dots,T, c \in \mathrm{C}, \tag{2}$$

Product inventory constraint

$$I_{p1} - B_{p1} = I_{p0} + \sum_{a\in \mathrm{A}} \rho_{ap} X_{a1} - d_{p1},$$
$$I_{pt} - B_{pt} = I_{pt-1} - B_{pt-1} + \sum_{a\in \mathrm{A}} \rho_{ap} X_{at} - d_{pt}, \qquad t = 2,\dots,T, p \in \mathrm{P}, \tag{3}$$

Production capacity constraint

$$\sum_{a\in \mathrm{A}} \delta_{ar} X_{at} \le M_{rt}, \qquad t = 1,\dots,T, r \in \mathrm{R}, \tag{4}$$

Non-negativity of all variables

$$X_{at} \ge 0, I_{ct} \ge 0, I_{pt} \ge 0, B_{pt} \ge 0, \qquad t = 1,\dots,T, p \in \mathrm{P}, c \in \mathrm{C}, a \in A. \tag{5}$$

The objective function (1) is a linear cost minimization equation. It consists of total inventory and backorder costs for all products and raw material cost for all material classes in the planning horizon. Constraint (2) ensures that the total inventory of raw material of class c at the end of period t is equal to its inventory in the previous period plus the quantity of material of class c supplied at the beginning of that period (s_{ct}) minus its total consumption in that period. Constraint (3) ensures that the sum of inventory (or backorder) of product p at the end of period t is equal to its inventory (or backorder) in the previous period plus the total production of that product in that period, minus the product demand for that period. Finally, constraint (4) requires that the total production do not exceed the available production capacity.

2.2. The Two-stage stochastic model with recourse for sawmill production planning with random yield

To include the random nature of processes yields in sawmill production planning, we expand the model (1)-(5) to a two-stage stochastic linear program with recourse. It is assumed that the random processes yields are modeled as scenarios with known probability distributions. We represent the random yield vector by $\underset{\sim}{\xi}$, where $\underset{\sim}{\xi} = \{\rho_{ap} \mid a \in \mathrm{A}, p \in \mathrm{P}\}$. We also represent each realization (scenario) of random processes yields by $\rho_{ap}(\underset{\sim}{\xi})$. We denote the total number of yield scenarios by N, and the probability of each scenario i by p^i, respectively. It should be emphasized that the stages of the two-stage stochastic problem do not refer to time periods. They correspond to steps in the decision making. In the first-stage (planning stage), the decision maker does not have any information about the processes yields due to lack of complete information on the characteristic of raw materials. However, the production plan should be determined before the complete information is available. Thus the first-stage decision variable is the production plan. In the second stage (plan implementation stage) when the realized yields are available, based on the first-stage decision, the recourse actions (inventory or backorder sizes) can be computed. The objective of the two-stage stochastic program with recourse would be to minimize the material consumption cost, plus the expected inventory and backorder costs (recourse costs) for all yield scenarios. The resulting deterministic equivalent formulation for the two-stage stochastic model is as follows:

$$\text{Minimize}\, Z = \sum_{c\in \mathrm{C}}\sum_{t=1}^{T}\sum_{a\in A} m_{ct}\phi_{ac}X_{at} + \sum_{i=1}^{N}\sum_{p\in \mathrm{P}}\sum_{t=1}^{T} p^{i}[h_{pt}I_{pt}^{i} + b_{pt}B_{pt}^{i}] \tag{6}$$

Subject to

$$I_{ct} = I_{ct-1} + s_{ct} - \sum_{a\in A}\phi_{ac}X_{at}, \qquad t=1,\dots,T, c\in \mathrm{C}, \tag{7}$$

$$I_{p1}^{i} - B_{p1}^{i} = I_{p0} + \sum_{a\in \mathrm{A}}\rho_{ap}(\underset{\sim}{\xi}^{i})X_{a1} - d_{p1},$$

$$I_{pt}^{i} - B_{pt}^{i} = I_{pt-1}^{i} - B_{pt-1}^{i} + \sum_{a\in \mathrm{A}}\rho_{ap}(\underset{\sim}{\xi}^{i})X_{at} - d_{pt}, \quad t=2,\dots,T, p\in \mathrm{P}, i=1,\dots,N, \tag{8}$$

$$\sum_{a\in \mathrm{A}}\delta_{ar}X_{at} \le M_{rt}, \qquad t=1,2,\dots,T, r\in \mathrm{R}, \tag{9}$$

$$X_{at}\ge 0, I_{ct}\ge 0, I_{pt}^{i}\ge 0, B_{pt}^{i}\ge 0, \qquad c\in C, p\in P, t=1,\dots,T, a\in A, i=1,\dots,N. \tag{10}$$

In the two-stage stochastic program (6)-(10), I_{pt}^{i} and B_{pt}^{i} denote the inventory and backorder sizes of product p in period t under yield scenario i, respectively.

3. Modeling the random processes yields in sawmills

To apply the proposed two-stage stochastic model (6)-(10) for sawmill production planning, as the first step, we should generate the scenarios for random processes yields. A scenario for the yields of process (a) (combination of a log class (c) and a cutting pattern (s)) in a sawmill is defined as possible quantities of lumbers that can be produced by cutting pattern (s) after sawing each log of class (c). As an example of the uncertain yields in sawmills, consider the cutting pattern (s) that can produce 6 products (P1, P2, P3, P4, P5, P6) after sawing the logs of class (*c*). Table 1 represents four scenarios among all possible scenarios for the uncertain yields of this process.

Table 1. Scenarios for yields of a process in sawmills

Scenarios	Products					
	P1	P2	P3	P4	P5	P6
1	1	0	1	0	1	1
2	2	1	1	0	1	0
3	1	0	0	1	1	1
4	2	0	0	1	0	1

In this work, we assume that all the logs that will be processed in the next planning horizon are supplied from the same discrete of forest. Hence, a stationary probability distribution can be considered for the quality of logs and uncertain processes yields during the planning horizon. We propose to consider the average yield of a random sample of logs in each log class as a scenario and to estimate the probability distribution for the average yields. Such scenarios with their probability distribution in sawmills can be determined as follows.

1) Take a sample of logs in each log class (e.g. 3000) and let them be processed by each cutting pattern. Compute the average yield for the sample.

2) Repeat step 1 for a number of replications (e.g. 30).

3) By the Central Limit Theorem (CLT) in statistics, the average yield has a normal distribution. Thus, based on the average yields computed for each replication in step 2, estimate the mean and variance of normal distribution corresponding to the average yield of each process.

It should be noted that, the implementation of step 1 in this approach is very difficult in sawmills. In fact, the high production speed in sawmills makes it almost impossible to track the logs through the line and to observe the result of sawing individual logs. As a more feasible alternative, we propose to use the set of yield scenarios generated by a log sawing simulator (Optitek). “Optitek” was developed by a research company for Canada's solid wood products industry (Forintek Canada Corp.). It was developed based on the characteristics of a large sample of logs in different log classes, as well as sawing rules available in Quebec sawmills. The inputs to this simulator include log class, cutting pattern, and the number of logs to be processed. The simulator considers the logs in the requested class with random physical and internal characteristics; afterwards it generates different quantity of lumbers (yields) for each log based on the sawing rules of the requested cutting pattern. Thus, in order to implement step 1 in the proposed scenario generation approach, a sample (e.g. 3000) of yields can be randomly taken among the set of scenarios already generated by Optitek, and the average yield for the sample can be computed.

4. Solution strategy

The two-stage stochastic model (6)-(10) can be solved by the linear programming solvers, namely CPLEX LP solver. However, regarding to the variety of characteristics in each log class in sawmills, a huge number of scenarios for processes yields can be expected. Thus, solving this model would be far beyond the present computational capacities. We can however use Monte Carlo sampling techniques, which consider only randomly, selected subsets of the set $\{\underset{\sim}{\xi}^1, \underset{\sim}{\xi}^2, \ldots, \underset{\sim}{\xi}^N\}$ to obtain approximate solutions. The sample average approximation (SAA) scheme (cf. Shapiro and Hommem-de-Mello, 1998, 2000; Mak et al., 1999), is selected as the solution approach in this work, which is described as follows.

Sample average approximation (SAA) scheme

In the SAA scheme, a random sample of *n* scenarios of the random vector $\underset{\sim}{\xi}$ is generated and the expectation $\sum_{i=1}^{N}\sum_{p\in \mathrm{P}}\sum_{t=1}^{T} p^i[h_{pt}I_{pt}^i + b_{pt}B_{pt}^i]$ is approximated by the sample average function $\frac{1}{n}\sum_{i=1}^{n}\sum_{p\in \mathrm{P}}\sum_{t=1}^{T}[h_{pt}I_{pt}^i + b_{pt}B_{pt}^i]$. In other words, the "true" problem (6)-(10) is approximated by the sample average approximation (SAA) problem (11).

$$\text{Minimize } \hat{Z} = \sum_{c\in \mathrm{C}}\sum_{t=1}^{T}\sum_{a\in A} m_{ct}\phi_{ac}X_{at} + \frac{1}{n}\sum_{i=1}^{n}\sum_{p\in \mathrm{P}}\sum_{t=1}^{T}[h_{pt}I_{pt}^i + b_{pt}B_{pt}^i] \tag{11}$$

Subject to

Constraints (7)-(10).

It can be shown that under mild regularity conditions, as the sample size *n* increases, the optimal solution vector $\hat{X}_n$ and optimal value $\hat{Z}_n$ of the SAA problem (11) converge with probability one to their true counterparts, and moreover $\hat{X}_n$ converges to an optimal solution of the true problem with probability approaching one exponentially fast (Shapiro and Hommem-de-Mello., 1998 and 2000). This convergence analysis suggests that a fairly good approximate solution to the true problem (6)-(10)

can be obtained by solving an SAA problem (11) with a modest sample size. The mentioned regularity conditions include: 1) the objective function of the stochastic model has finite mean and variance, 2) the independent identically distributed (i.i.d.) observations of vector $\underset{\sim}{\xi}$ can be generated, 3) instances of SAA problem can be solved for sufficiently large *n* to generate "good" bounding information, and 4) the objective function of the stochastic model can be evaluated exactly for specific values of X_{at} and realizations of vector $\underset{\sim}{\xi}$. It is evident that the mentioned regularity conditions are satisfied for our problem, especially due to considering a normal distribution for the random yields. In practice, the SAA scheme involves repeated solutions of the SAA problem (11) with independent samples. Statistical confidence intervals are then derived on the quality of the approximate solutions (Mak et al., 1999). According to the work of Mak et al. (1999), an obvious approach to test solution quality for a candidate solution ($\bar{X}$) is to bound the optimality gap, defined as $E_{\underset{\sim}{\xi}}[f(\bar{X},\underset{\sim}{\xi})]-z^{*}$ using standard statistical procedures, where $f(\bar{X},\underset{\sim}{\xi})$ and z^{*} are the true objective value for $\bar{X}$ and the true optimal solution to the problem (6)-(10), respectively, and $E_{\underset{\sim}{\xi}}[f(\bar{X},\underset{\sim}{\xi})]$ is the expected value of $f(\bar{X},\underset{\sim}{\xi})$. In our work, a sampling procedure based on common random numbers (CRN) is used to construct the optimality gap confidence interval which provides significance variance reduction over naive sampling, as shown in (Mak et al., 1999). This approach is described next.

The SAA algorithm (with common random number streams)

Step 1- Generate n_g independent identically distributed (i.i.d.) batches of samples each of size *n* from the distribution of $\underset{\sim}{\xi}$, i.e., $\left\{\underset{\sim}{\xi}_j^{1},\underset{\sim}{\xi}_j^{2},\ldots,\underset{\sim}{\xi}_j^{n}\right\}$ for $j=1,\ldots,n_g$. For each sample, solve the corresponding SAA problem (11). Let $\hat{Z}_n^j$ and $\hat{X}_n^j$, $j=1,\ldots,n_g$, be the corresponding optimal objective value and an optimal solution, respectively.

Step 2- Compute

$$\bar{Z}_{n,n_g}=\frac{1}{n_g}\sum_{j=1}^{n_g}\hat{Z}_n^j, \text{ and} \qquad (12)$$

$$s^2_{\bar{Z}_{n,n_g}}=\frac{1}{n_g(n_g-1)}\sum_{j=1}^{n_g}(\hat{Z}_{n_g}^j-\bar{Z}_{n,n_g})^2. \qquad (13)$$

It is well known that the expected value of $\hat{Z}_n$ is less than or equal to the optimal value z^* of the true problem (see e.g., Mak et al., 1999). Since $\bar{Z}_{n,n_g}$ is an unbiased estimator of $E[\hat{Z}_n]$, we obtain that $E[\bar{Z}_{n,n_g}] \le z^*$. Thus $\bar{Z}_{n,n_g}$ provides a lower statistical bound for the optimal value z^* of the true problem (6)-(10) and $s^2_{\bar{Z}_{n,n_g}}$ is an estimate of the variance of this estimator.

Step 3- Choose a candidate feasible solution $\bar{X}$ of the true problem, for example, a computed $\hat{X}^j_{n'}$ by using a sample size (n') larger than used for lower bound estimation (n). Estimate the true objective function value $f(\bar{X})$ for all batches of samples ($j=1,\dots,n_g$) as follows.

$$\text{Minimize } \tilde{f}^j_n(\bar{X}) = \sum_{c\in C}\sum_{t=1}^{T}\sum_{a\in A} m_{ct}\phi_{ac}\bar{X} + \frac{1}{n}\sum_{i=1}^{n}\sum_{p\in P}\sum_{t=1}^{T}[h_{pt}I^i_{pt} + b_{pt}B^i_{pt}]. \quad (14)$$

Subject to

Constraints (7)-(10).

Step 4- Compute the observations of the optimality gap G^j_n for the candidate solution $\bar{X}$ for all $j=1,\dots,n_g$ as follows.

$$G^j_n = \tilde{f}^j_n(\bar{X}) - \hat{Z}^j_n. \quad (15)$$

It has been shown in Mak et al. (1999) that:

$$E\underbrace{\left[\tilde{f}_n(\bar{X}) - \hat{Z}_n\right]}_{G_n} \ge E_{\underset{\sim}{\xi}}[f(\bar{X},\underset{\sim}{\xi})] - z^*. \quad (16)$$

where $f(\bar{X},\underset{\sim}{\xi})$ and z^* are the true objective value for $\bar{X}$ and the true optimal solution to the problem (6)-(10), respectively, and ($E_{\underset{\sim}{\xi}}[f(\bar{X},\underset{\sim}{\xi})] - z^*$) is the true optimality gap for the candidate solution $\bar{X}$. We also have:

$$\sqrt{n_g}\left[\bar{G}_{n_g} - EG_n\right] \Rightarrow \mathrm{N}(0,\sigma^2_g) \quad \text{as } n_g \to \infty$$

where $\sigma^2_g = \text{var } G_n$.

Step 5- Compute the sample mean and sample variance for the optimality gap G^j_n as follows.

$$\bar{G}_{n_g} = \frac{1}{n_g}\sum_{j=1}^{n_g} G_n^j, \text{and} \tag{17}$$

$$s_{G_n^j}^2 = \frac{1}{n_g(n_g-1)}\sum_{j=1}^{n_g}(G_n^j - \bar{G}_{n_g})^2.$$

Step 6- Compute the approximate $(1-\alpha)$-level confidence interval for the optimality gap of $\bar{X}$ as

$$\left[((\bar{G}_{n_g} - \tilde{\varepsilon}_g) \vee 0), (\bar{G}_{n_g} + \tilde{\varepsilon}_g)\right], \text{ where } \tilde{\varepsilon}_g = \frac{t_{n_g-1,\alpha} s_{G_n^j}}{\sqrt{n_g}} \quad . \tag{18}$$

5. Validation of the stochastic sawmill production planning model by Monte Carlo simulation

In this section, we compare the plans proposed by the stochastic and deterministic sawmill production planning models. As we mentioned before, we assume that the company is very service sensitive, i.e., the realized total backorder size after implementation of production plan is more crucial than the realized inventory size. Thus, the following key indicators of performance are considered to compare the deterministic and stochastic models:

1) Backorder gap (BO gap): the gap between the realized total backorder size of the deterministic and the stochastic models' plans, after implementing the mentioned plans.

2) Plan precision: the gap between the planned total backorder size determined by the production planning model and the realized total backorder size, after implementing the model's plan. This indicator evaluates also the extent to which the yield scenarios considered in the stochastic model are close to the scenarios that can be observed in the real production process.

In order to compute the total backorder size after implementing the plans, we propose to use Monte Carlo simulation. The main objective of this simulation is to implement the production plans virtually, by considering the yield scenarios that might be realized during the plan implementation in the realistic scale sawmills. Hence, the following features are considered for the simulator:

1) To get the production plans proposed by the deterministic and stochastic models as well as the products demand, as the inputs.

2) To simulate the production plan implementation based on the received production plan as follows:

2.1) To determine a sample size equal to the number of times each process should be run in each period (production plan).

2.2) To generate randomly a sample of scenarios (with the size determined in 2.1) for the yields of each process, from the set of possible scenarios for the yields of that process. It should be mentioned that this step is equivalent to select a random sample of logs in each class to be sawn by each cutting pattern, while implementing the production plan in sawmills.

3) To compute the total production size of each product at the end of each period, after simulating the plan implementation for that period (step 2).

4) To compute the backorder or inventory size of each product in each period based on the total production size of that product (computed in step 3) and its demand for that period.

Figure 1 illustrates the main features of the simulator which is designed to simulate the plans implementation in sawmills.

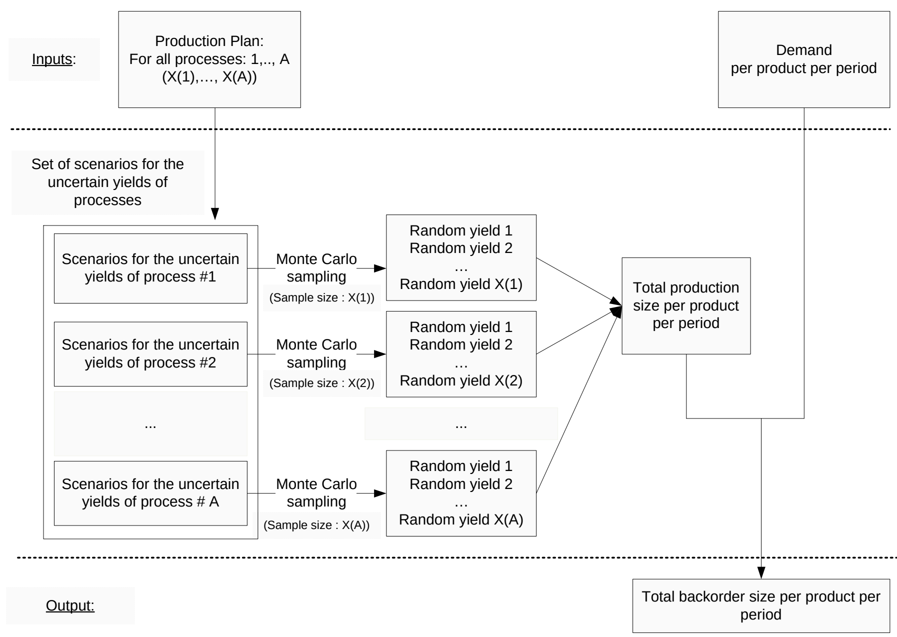

Figure 1. Simulation of the production plans implementation in sawmills

6. Computational results

In this section, we describe the numerical experiments using the proposed two-stage stochastic model to address a medium capacity sawmill production planning problem. We first describe the characteristics of the test industrial problem and some implementation details; then, we comment on the quality of the stochastic model solutions determined by the SAA scheme; finally, we compare the stochastic and mean-value deterministic models' plans by the proposed Monte Carlo simulation approach (see section 5), for different demand levels.

6.1. Data and implementation

The proposed two-stage stochastic program with recourse in this paper is applied for a prototype sawmill. The prototype sawmill is a typical medium capacity softwood sawmill located in Quebec (Canada). The sawmill focuses on sawing high-grade products to the domestic markets as well as export products to the USA. It is assumed that the input bucked logs into the sawmill are categorized into 3 classes. 5 different cutting patterns are available. The sawmill produces 27 products of custom sizes (e.g. 2(in)$\times$4(in), 2(in)$\times$6(in) lumbers) in four lengths. In other words, there are 15 processes able to produce 27 products with random yields. We consider two bottleneck machines: Trimmer and Bull. The planning horizon consists of 30 periods (days). Product demand in each period is assumed to be deterministic which is determined based on the received orders. Lumbers that remain from one period to the next are subject to a unit holding cost. The unsatisfied demand is penalized by a unit backorder cost. We assume that the company is very service sensitive and wishes to fulfill customer demands on time as much as possible. Hence, the inventory costs of products are considered much lower than their backorder costs. The inventory holding cost is computed by multiplying the interest rate (per period) by the lumber price; the lumber price is considered as the backorder cost. It would be worth mentioning that the data used in this example are based on the gathered data from different sawmills in Quebec province (Canada). As the list of custom sizes, machine parameters and prices are proprietary; they are not reported in this paper.

Recall from section 4 that the SAA method calls for the solution of n_g instances of the approximating stochastic program (11), each involving n sampled scenarios. Statistical validation of a candidate solution is then carried out by evaluating the objective function value using the same n sampled scenarios in each batch. In our implementation test, we used n=30 and 100; and $n_g = 30$. Our candidate solutions are computed by solving the SAA problem (11) with $n' = 100$ and 150. To illustrate the complexity of solving (11) within the SAA scheme, we present the sizes of the deterministic equivalents of the SAA problems corresponding to the different values of n in Table 2.

Table 2. Size of the deterministic equivalent of the SAA problem

n	Number of constraints	Number of variables
1	960	2160
30	24450	49140
100	81150	162540
150	121650	243540

The SAA scheme was implemented in OPL Studio 3.7.1. CPLEX 9 LP solver is used for solving the deterministic equivalents for different instances of SAA problems as well as for calculating the true objective function value for the candidate solutions. The simulator is programmed by Java. All computations were carried out on a Pentium(R) IV 1.8 GHz PC with 512 MB RAM running Windows XP.

6.2. Quality of the stochastic solutions

In this section, we present the results of applying the SAA scheme for our test problem and the evaluation of quality of candidate approximate solutions. Point estimates (see (12) and (13)) of the lower statistical bound for the optimal value of the stochastic problem are reported in table 3. They are computed based on 30 batches of sampled scenarios with 2 different batch sizes. Table 4 displays the quality of 2 candidate solutions and contains the 95% confidence intervals on their optimality gaps based on CRN method (see section 4). The candidate solutions $\bar{X}^{100}, \bar{X}^{150}$ for the CRN strategy are computed by solving the approximating problem (11) that includes 100 and 150 scenarios. The CPU times for computing each candidate solution are also reported in table 4.

Table 3. Lower bound estimation results for the optimal value (n_g = 30 batches)

Batch size (n)	30	100
Average ($\bar{Z}_{n,n_g}$)	1,923,901	1,924,380
SD ($s_{\bar{Z}_{n,n_g}}$)	4,730	4,068

Table 4. Optimality gaps for candidate solutions

Candidate solution	$\bar{X}^{100}$	$\bar{X}^{150}$
Batch size (n)	30	100
No. of batches (n_g)	30	30
Point estimate ($\bar{G}_{n_g}$)	1710	918
Error estimate ($\alpha = 95\%$) ($\tilde{\varepsilon}_g$)	737	284
Confidence interval (95%)	[0 , 2447]	[0 , 1202]
CPU time (minutes)	20	25

As it can be observed from Table 4, by increasing the sample size, the quality of approximate solutions improves and the tighter confidence intervals for the optimality gaps of candidate solutions are constructed. Finally we can conclude that, by considering a moderate number of scenarios (150 scenarios) among the potential enormous number of scenarios, we obtain an approximate solution in a reasonable amount of time with an optimality gap of [0, 1202] which is about 0.006% of the lower bound of the real optimal value (see Tables 3). Thus, this solution can be accepted as a good approximation to the optimal solution of the original stochastic model (6)-(10).

6.3. Comparison between the stochastic and deterministic sawmill production planning models

In this section, the results of comparison between the two-stage stochastic and mean-value deterministic sawmill production planning models, through Monte Carlo simulation (see section 5), are provided. The comparison is carried out for the sawmill example described in 6.1. Four different demand levels (D1, D2, D3, D4) are considered, where D2=2$\times$D1, D3=3$\times$D1, D4=4$\times$D1. For each demand level, 60 demand scenarios are generated randomly which are distinguished by the distribution of total demand between different products. Hence, a total of 240 (4$\times$60) test problems are solved by the deterministic LP and stochastic models. The simulation of implementing the production plans is run for 1000 replications. The expected total backorder size computed in 1000 replications is used to compute the key indicators of performances (see section 5) for the test problems. Table 5 includes the mean and standard deviation (SD) of the backorder gap (BO gap) as well as the plan precision computed for the 60 test problems, corresponding to the 60 demand scenarios, in each of the 4 demand levels. It would be worth mentioning that the values of BO gap and plan precision presented in table 5 are computed as follows:

BO gap = $100\times(BO_D - BO_s)/BO_D$

Plan precision = $100\times(BO_{sim} - BO_{plan})/BO_{sim}$

where

BO_D : The expected realized total backorder size of the deterministic model after plan implementation (computed through Monte Carlo simulation)

BO_S : The expected realized total backorder size of the stochastic model after plan implementation (computed through Monte Carlo simulation)

BO_{sim} : The expected realized total backorder size after plan implementation (computed through Monte Carlo simulation)

BO_{plan} : Total backorder size determined by the production planning model

To illustrate how the results in table 5 can be interpreted, the following examples are provided. The mean and the standard deviation (SD) of BO gap for the demand level D1, determined as 75% and 21%, respectively, in table 5 can be interpreted as follows: the realized total backorder size of stochastic model plan is 54% to 96% smaller than the realized total backorder size of deterministic model plan. Thus, it can be concluded that the stochastic model outperforms the deterministic model in proposing the production plans with lower total backorder size.

Table 5. Comparison between the deterministic and stochastic sawmill production planning models

Sawmill production planning model	D1				D2				D3				D4			
	BO gap		Plan precision		BO gap		Plan precision		BO gap		Plan precision		BO gap		Plan precision	
	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)	Mean (%)	SD (%)
Deterministic	-	-	145	350	-	-	35	138	-	-	7	93	-	-	-13	40
Two-stage Stochastic	75	21	-40	19	58	33	-34	15	39	25	-26	16	31	21	-19	10

The mean and standard deviation of plan precision for the demand level D3 in the stochastic model, indicated as -26% and 16%, respectively, can be interpreted as follows: The realized total backorder size of production plan proposed by the stochastic model is 10% to 26% smaller than the total backorder size that was initially determined by this model. A positive value for plan precision indicates that the realized total backorder size through Monte Carlo simulation is larger than the planned total backorder size, as in the case of deterministic model for demand levels D1, D2 and D3.

We next analyze the results provided in table 5 to compare the performance of the deterministic and the stochastic sawmill production planning models in terms of their total backorder size as well as their plan precision. Figure 2 compares the mean backorder gap (BO gap) between the stochastic and the deterministic models, for the four demand levels. As it can be observed in table 5 and figure 2, the production plan proposed by the stochastic model results smaller realized total backorder size (after implementing the plan) than the deterministic model plan, for the four demand levels. However, the gap between the total backorder size of the stochastic and the deterministic models' plans decreases, as the demand is increased. This should make no surprise. As we mentioned before, the sawmill example is a medium capacity sawmill where thousands of logs are sawn in each period in the planning horizon. By the law of large numbers (LLN) in statistics, as the demand is increased, the average yields of each process in each period, observed through Monte Carlo sampling in the plan implementation simulator, will be closer to their expected values which are considered in the deterministic model.

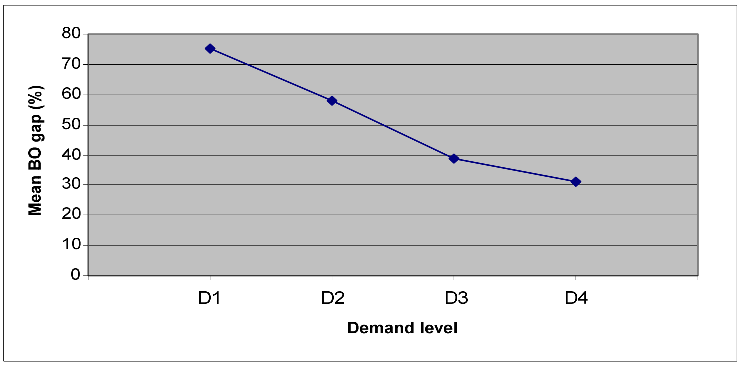

Figure 2. Mean backorder gap (BO gap) of the stochastic and deterministic sawmill production planning models

Figure 3 compares the mean of plan precision of the stochastic and the deterministic model, for the four demand levels. As it can be observed in table 5 and figure 3, the precision of production plan proposed by the stochastic model is higher than the deterministic one, for the four demand levels. As the demand increases, the average yields observed after implementing the plan through Monte Carlo simulation, get close to the average yield scenarios considered in the stochastic model. Hence the precision of plans of the stochastic model improves for the larger volumes of demand. For the lower demand levels, the stochastic model proposes relatively pessimistic plans. On the other hand, the deterministic model provides the optimistic plans for demand levels D1, D2 and D3, since it does not take into account different scenarios for random yield. However, as the demand increases, the average yields of each process in each period observed through Monte Carlo simulation get closer to their expected values which are used in the deterministic model (LLN). Thus, the precision of deterministic model plan increases, as the demand is increased.

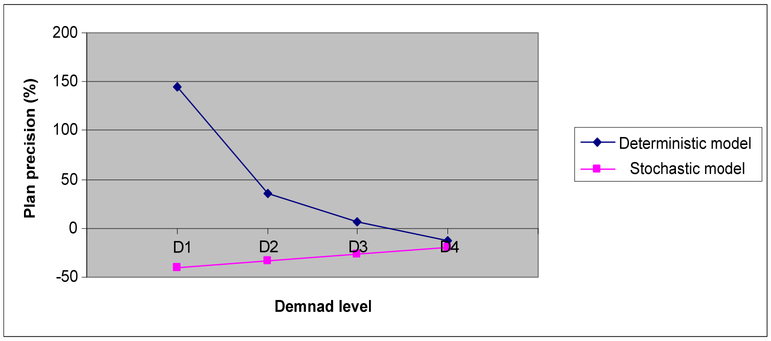

Figure 3. Mean plan precision comparison of deterministic and stochastic sawmill production planning models

Regarding the above comparisons, it is clear that the two-stage stochastic model provides more realistic production plans in sawmills, in terms of the realized backorder size, than the mean value deterministic LP model. The deterministic model provides the optimistic plans, since it considers the deterministic yields (expected values). On the other hand, as the stochastic model considers different scenarios for random yields and finds a production plan with minimum expected backorder and inventory size for all the yield scenarios, the production plans provided by this model are more realistic.

7. Conclusions

In this paper, we developed a two-stage stochastic programming model for multi-period sawmill production planning under the uncertainty of processes yields. The SAA method was implemented to solve the stochastic model which provided us an efficient framework for identifying and statistically testing a variety of candidate production plans. We also proposed a validation approach to compare the plans proposed by the stochastic and deterministic sawmill production planning models, which is based on Monte Carlo simulation. We provided the empirical results for production planning in a medium capacity prototype sawmill and we identified the candidate plans in a reasonable amount of time by solving the approximate SAA problem. Furthermore, the confidence intervals for the optimality gap of candidate solutions were constructed by common random number (CRN) streams. The comparison between the two-stage stochastic and deterministic sawmill production planning models was carried out for 4 demand levels. Our results revealed that the production plans proposed by the stochastic model are considerably superior to those obtained by traditional mean-value deterministic model. Although these results are found for sawmill production planning, the proposed approach in this work can be applied for production planning in other manufacturing environments where non-homogeneous and random characteristics of raw materials result random yield. Future research will consider the products demands as random variables in order to obtain more realistic production plans.

8. Acknowledgments

The authors would like to acknowledge the financial support provided by the Forest E-business Research Consortium (FOR@C) of Université Laval, and would like to thank specially, Jonathan Gaudreault, Philippe Marier, Sébastien Lemieux, and Christian Rouleau, for their technical support.

9. References

Bakir, M.A. and Byrune, M.D. (1998) Stochastic linear optimization of an MPMP production planning model. *International Journal of Production Economics*, 55, 87-96.

Birge, J.R. and Louveaux, F. (1997) *Introduction to stochastic programming*. Springer, New York.

Escudero, L.F., Kamesam, P.V., King, A.J. and Wets, R.J-B. (1993) Production planning via scenario modeling. *Annals of Operations Research*, 43, 311-335.

Gaudreault, J., Rousseau, A., Frayret, J.M. and Cid, F. (2004) Planification opérationnelle du sciage. *For@c technical document*, Quebec (Canada).

Johnson, L.A. and Montgomery D.C. (1974) *Operations research in production planning, scheduling, and inventory control*. John Wiley & Sons, New York.

Kall, P. and Wallace, S.W. (1994) *Stochastic programming*. John Wiley & sons, New York.

Kall, P. and Mayer J. (2005) *Stochastic linear programming*. Springer's International Series, New York.

Kazemi Zanjani, M., Nourelfath, M. and Ait-Kadi, D. (2007) Production planning for simultaneous production of multiple products by alternative processes with random yield, in *Proceedings of 7e Congrès international de génie industriel (2007),* Trois-Rivières (Canada).

Mak, W.K., Morton, D.P. and Wood R.K. (1999) Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations research letters*, 24, 47-56.

Maness, T.C. and Adams, D.M. (1991) The combined optimization of log bucking and sawing strategies. *Wood and Fiber Science*, 23, 296-314.

Maness, T.C. and Norton, S.E. (2002) Multiple-period combined optimization approach to forest production planning. *Scandinavian Journal of Forest Research*, 17, 460-471.

Mendoza, G.A., Meimban, R.J., Luppold, W.J. and Arman P.A. (1991) Combined log inventory and process simulation models for the planning and control of sawmill operations, in *Proceedings of the 23rd GIRP International Seminar on Manufacturing Systems,* Nancy (France).

Shapiro A. and Hommem-de-Mello T. (1998) A simulation based approach to two-stage stochastic programming with recourse. *Mathematical Programming*, 81, 301-325.

Shapiro A. and Hommem-de-Mello T. (2000) On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM Journal on Optimization*, 11(1), 70-86.

Sipper D. and Bulfin R. (1997) *Production: planning, control, and integration*. McGraw-Hill, New York.

Sox, CH.R. and Muckstadt, J. A. (1996) Multi-item, multi-period production planning with uncertain demand. *IIE Transactions*, 28(11), 891-900.