Robust Scheduling in Forest Operations Planning

by

Lui Cheng, Lim

B.Eng. Mechanical Engineering (2007)
National University of Singapore

Submitted to the School of Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Computation for Design and Optimization

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2008

Author ....
School of Engineering
August 11, 2008

Certified by ....
Jorge R. Vera
Visiting Associate Professor of Sloan School of Management
Thesis Supervisor

Accepted by ....
Jaime Peraire
Professor of Aeronautics and Astronautics
Codirector, Computations for Design and Optimization

Robust Scheduling in Forest Operations Planning

by

Lui Cheng, Lim

Submitted to the School of Engineering
on August 11, 2008, in partial fulfillment of the
requirements for the degree of
Master of Science in Computation for Design and Optimization

Abstract

Forest operations planning is a complex decision process which considers multiple objectives on the strategic, tactical and operational horizons. Decisions such as where to harvest and in what order over different time periods are just some of the many diverse and complex decisions that are needed to be made. An important issue in real-world optimization of forest harvesting planning is how to treat uncertainty of a biological nature, namely the uncertainty due to different growth rates of trees which affects their respective yields. Another important issue is in the effective use of high capital intensive forest harvesting machinery by suitable routing and scheduling assignments. The focus of this thesis is to investigate the effects of incorporating the robust formulation and a machinery assignment problem collectively to a forest harvesting model.

The amount of variability in the harvest yield can be measured by sampling from historical data and suitable protection against uncertainty can be set after incorporating the use of a suitable robust formulation. A trade off between robustness to uncertainty with the deterioration in the objective value ensues. Using models based on industrial and slightly modified data, both the robust and routing formulations have been shown to affect the solution and its underlying structure thus making them necessary considerations. A study of feasibility using Monte Carlo simulation is then undertaken to evaluate the difference in average performances of the formulations as well as to obtain a method of setting the required protections with an acceptable probability of infeasibility under a given set of scenarios.

Thesis Supervisor: Jorge R. Vera
Title: Visiting Associate Professor of Sloan School of Management

Acknowledgments

I would like to thank the following people for their help and guidance in the course of this thesis research.

First and foremost, I would like to express my gratitude and appreciation for my thesis advisor, Jorge Vera, for being a great teacher and friend. His patience in guiding me as well as his willingness in exploring my ideas has been crucial to my research. His sincere, humble and open nature are lessons in real life about how to conduct myself in person and his mental acuity challenges me on several fronts. It is my great pleasure to have him as my supervisor.

I am also very thankful to the Singapore-MIT Alliance for giving me the opportunity to study in MIT through their fellowship funding. Special thanks goes to Jocelyn and John, for their efforts in creating invaluable memories for us in MIT.

My gratefulness also goes to my fellow colleagues from the Department of Computation for Design and Optimization with whom I have gone through the same grueling courses and fun-filled activities. They are like a family to me. I would like to specially thank Ting Ting for her care and concern during this entire phrase.

Last but not least, I must also thank my family and friends for their love and support throughout my stay in MIT.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Thesis Background and Motivation

Forest harvesting is an activity that has progressed from harvesting in small parcels of land for building materials and firewood in the past, to becoming a large scale commercial activity over the past century. Efficient forest management involves making a variety of inter-related decisions to achieve goals set for the operational, tactical and strategic horizons. At the strategic level, forest decisions have to be made for road building, fleet management and sustainability. In the past few years, various other environmental considerations such as biodiversity, wildlife protection and global warming control are also included when making decisions at the strategic level. At the tactical level, annual harvest plans, road upgrades and equipment utilization decisions are made. Lastly, at the operational level, harvest scheduling, bucking, truck dispatching are just some of the daily decisions made for smooth and cost effective operations. The interested reader is directed to [11] for a more thorough discussion on the wide ranging decision making problems in the forest industry.

The use of optimization models to aid such decision making processes in the forest industry is therefore required and has been acknowledged in awards such as the Edelman prize in 1998 [10]. In many supply chain management models, market uncertainty from both the demand and supply sides are important factors that needs to be taken into consideration before making critical decisions. Collected field data can be subjected to measurement errors, which leads to data uncertainty in the optimization models. From the demand side, demand is mostly projected and seldom known exactly, and optimal prices are estimated as they depend on the evolution of the commodities market. From the supply side, uncertainties may stem from incomplete knowledge of the amount and quality of available resources as in the case of oil drilling and even biological factors in agricultural and forest industries. Such uncertainties can be addressed using a variety of techniques such as stochastic programming, dynamic programming, chance-constrained optimization and robust optimization.

Robust optimization is a methodology in optimization that addresses the need to hedge against uncertainty that may arise from various sources. Typically, a data parameter subjected to noise can be modeled as a distributed random variable. However, full knowledge of the distributions of the random variables are seldomly known and thus robust optimization deals well in this aspect as only basic knowledge of the random variables is enough to define the model. Robust optimization then protects for variations within the deviation of these random variables. Robust optimizaton has been sucessfully applied in various field including engineering, finance and supply chain optimization. This thesis implements robust optimization to mitigate the risk of not meeting forest product demands which may have contractual or other penalties and a more detailed discussion can be found in Chapter 2.

The other part of the problem examined in this research deals with the efficient use of harvesting machinery. Commonly, 3 types of harvesting machines are used in harvesting a forest: Skidders, Harvesters and Towers.

The 3 types of machines are shown in Figures 1-1, 1-2, and 1-3 respectively. A skidder is similar to a 4 wheel drive tractor that is used for pulling logged trees in a process called "skidding" where the logs are transported from the original site to a landing site for further transportation to a production facility. A harvester is another type

Figure 1-1: Example of a Skidder

Figure 1-2: Example of a Harvester

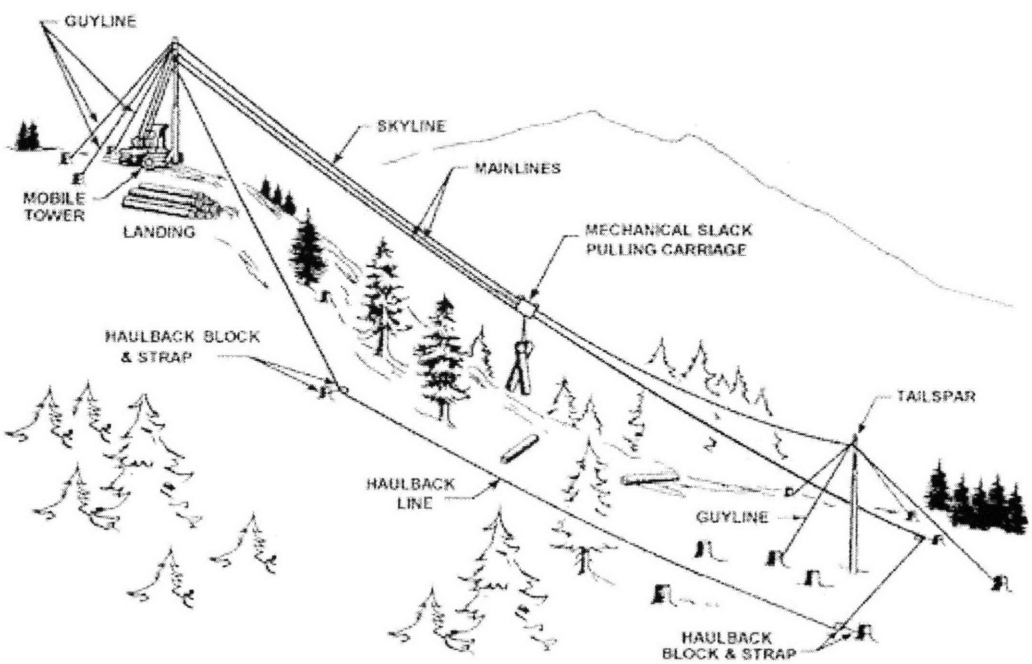

Figure 1-3: Tower with mechanical slack pulling carriage system

of heavy vehicle which can be used for felling and bucking trees and are employed effectively in level to moderately steep terrain. Lastly, towers are crane-like structures perched on a hilltop which drags the logs to level terrain via cables analogous to a ski lift and is suitable for harvesting on steep slopes. Therefore, depending on the terrain, different kinds of machinery are used. All these harvesting machines are high capital intensive machinery which incurs not only high fixed and operational costs but also high assignment costs when they have to be shifted from one forest parcel to another. In particular, the towers require extremely high setup costs and time due to the complexity of setting up the system as can be seen from Figure 1-3. Therefore, the efficient scheduling of these equipments to the right forests is an important problem which can result in significant cost savings if done correctly. Furthermore, the problem is further complicated by the fact that not all forest parcels are equal neither in size nor in their composition. Each forest parcel comprises of different species of trees that have different grades and which are used for making different products which affects the yield of each forest stand. Hence, the efficiency of each type of machinery is different within each forest parcel.

We are interested in the implementation of the robust formulation and machinery scheduling problem and to examine the differences in solutions obtained from a nominal forest problem, a robust model with protection from uncertainty and finally a robust model with machine scheduling considerations. The uncertainty considered is in the yield of the forest which is closely dependent on the tree species and results in different grades of wood.

1.2 Literature Review

Harvesting Planning with scheduling of machinery

In [14], Karlsson et al. addresses the problem of jointly determining the optimal decisions for a short-term harvesting planning problem and crew scheduling problem. They define a harvest crew as consisting of a harvester and a forwarder working in tandem. The planning horizon considered was 4-6 weeks with 6 harvest crews and an MIP problem was developed for his purposes. The solution approach used was by solving deterministically using a commercial solver, CPLEX and also by using a pseudo column generation approach where the required schedules were generated *a priori*. Their model presents some similarity with the harvesting and scheduling model that is used in this research. Certain aspects considered in his paper that are not pursued in this research include the age-related cost of storing and road network for transportation and maintenance. However, these are not considered in this research as attention is focused on uncertainty containment and the effects of a robust formulation on scheduling.

Other approaches for a machine location problem

In [19], Vera et al. considered the use of a Lagrangian relaxation approach for solving a large scale harvesting machine location problem. The work deals jointly with selecting the locations for the machinery as well as designing the access road network and is therefore the combination of 2 well known hard combinatorial optimization problems. The solution obtained using CPLEX after 600 minutes only yield a solution with a feasible gap of 27% testifying to the difficulty in solving the problem to optimality. Their approach work towards reducing this gap by decomposing the problem into the 2 subproblems using Lagrangian relaxation, strengthening the subproblems and solving using subgradient and subgradient-hybrid approaches. In a follow up research [15], Diaz et al. considered the use of a tabu search approach for solving the same problem. Their solution algorithm consists of a basic tabu search algorithm coupled with minimum spanning tree algorithms and Steiner tree heuristics to evaluate the neighbourhood of a solution at each point.

Robust optimization in forest operations planning

In [16], Maturana et al. addresses the use of robust optimization in forest operations planning and their work forms the backbone for this research. In this paper, they considered the use of robust optimization to handle the uncertainty of meeting demand due to biological factors. 3 solution approaches to solving the robust counterpart of a basic forest operations planning problem that considers harvesting and transportation to demand points were used and compared. The need for a comparison of 3 different approaches arose from the conflict in the objective and constraints after writing the robust counterpart and the first approach involved ignoring the robust counterpart of the objective in favour of the constraints and examining the resulting solutions. The second approach, termed adversarial approach, used an iterative solution approach developed by Bienstock [7] which allowed correlation in the data and therefore aided in managing the above-mentioned conflict. The last approach was a heuristic approach that used as part of the algorithm, one iteration of the adversarial approach. The paper concluded with a difference in the solutions resulting from the different approaches but the difference is limited and the adversarial approach might have resulted in a slightly conservative solution. More discussion follows in Section 3.2 and some relevant results are shown in section 4.2.

1.3 Thesis outline

In Chapter 2, the various robust formulations for an optimization model are discussed. Chapter 3 presents the nominal and robust forest harvesting model and constraints added to the model to extend the consideration to machinery scheduling are also discussed here. Chapter 4 focuses on the use of models using industrial and slightly modified data to exemplify the effects of the robust formulation and routing consideration to the harvesting decisions and strengthens the argument for their inclusion in models for forest harvesting planning. Results of a Monte Carlo simulation which investigates the average properties of the various models as well as gives an indication for the level of conservativeness to apply for subsequent harvesting models are presented in Chapter 5. Finally, in Chapter 6, we conclude the thesis with a summary and suggestions for future research directions.

Chapter 2

Robust Formulations

This chapter describes a brief summary of the robust formulations developed by Soyster, Ben-Tal and Nemirovski which gives us a basic insight to the development of robust optimization and a detailed approach developed by Bertsimas and Sim which was eventually used for this work.

The development of robust optimization was initiated by Soyster in 1973 [18] who proposed a linear optimization model to obtain a solution that is feasible for all data that belongs in a convex set. The model was deemed too conservative and traded off too much optimality for a robust solution. Subsequently, steps were taken by Ben-Tal and Nemirovski [1, 2, 4] and El-Ghaoui and Lebret [12, 13] who proposed models that involved solving the robust counterparts of a nominal problem by casting them in the form of conic quadratic problems. However, this resulted in nonlinear models which require significantly more computational resources than linear models. Bertsimas and Sim [5, 6] thereafter proposed an approach for robust linear optimization that is not only linear but also applicable to discrete problems and allows the user to vary the level of conservationism for every constraint.

Based on considerations for the application to the foresting industry, we believe that a few key issues will affect the choice of a suitable robust formulation to be used:

1. Conservativeness of the optimal solution
2. Computational tractability
3. Flexibility in varying the level of conservativeness

The nominal formulation of a typical linear programming problem is:

$$\begin{aligned} \text{maximise} \quad & c'x \\ \text{subject to} \quad & Ax \leq b \\ & l \leq x \leq u. \end{aligned}$$

Using this initial formulation, the assumption is that the data uncertainty only affects the elements of the matrix A. Without loss of generality, the cost function $c'x$ can be assumed to be without uncertainty as it can be easily replaced by another variable, z, and brought into the constraints using $z - c'x \leq 0$ which can be easily incorporated into the matrix inequality, $Ax \leq b$.

The model of uncertainty undertaken is as described by Ben-Tal and Nemirovski [3] whereby a row of matrix A is indexed using i with J_i representing the set of coefficients in row i that are subjected to uncertainty. Each uncertain element, $a_{ij}, j \in J_i$, is modeled as a symmetric and bounded random variable, $\tilde{a}_{ij}, j \in J_i$, that takes values within the interval $[a_{ij} - \tilde{a}_{ij}, a_{ij} + \tilde{a}_{ij}]$, where $\tilde{a}_{ij}$ is the amount of deviation of the element. The budget of uncertainty, Γ, is defined as the maximum amount of uncertainty acceptable within the data and is formulated as a constraint, $\sum_{j=1}^{n} \frac{1}{\tilde{a}_{ij}} |\tilde{a}_{ij} - a_{ij}| \leq \Gamma$ and therefore allows one to control the "price of robustness" as defined by Bertsimas and Sim [5] by varying the value of Γ.

When the value of Γ_i is equal to 0 then the problem is the original nominal model which does not take uncertainty into account and when Γ_i is equal to the cardinality of J_i, the problem is the equivalent to Soyster's formulation which offers full protection from uncertainty.

2.1 Soyster's full protection robust formulation

Soyster considers the matrix A as a series of columns $(a_1, a_2, \ldots, a_n)$ with $a_j, j = 1 \ldots n$, being the column vectors that are subjected to uncertainty. It is only known with certainty that the vector lies within a hypercube with center a_{ij} and half-length $\hat{a}_{ij}$ that represents the actual value and the maximum deviation respectively. By separating the actual vector and the deviation and allocating a decision variable, $y_j, j = 1, \ldots, n$, to the deviation, the robust formulation considered by Soyster can be represented as:

$$
\begin{aligned}
\text{maximise} \quad & c'x \\
\text{subject to} \quad & \sum_j a_{ij}x_j + \sum_{j \in J_i} \hat{a}_{ij}y_j \leq b_i && \forall i \\
& -y_j \leq x_j \leq y_j && \forall j \\
& l \leq x \leq u \\
& y \geq 0.
\end{aligned}
$$

It is noted that y_j is the absolute value of the original decision variable at optimality, $|x_j^*|$ and that the model can be easily shown to be feasible for all values of $\tilde{a}_{ij}, j \in J_i$. The 2nd term on the left hand side in the 1st constraint, $\sum_{j \in J_i} \hat{a}_{ij}|x_j|$, gives the required robustness of solution by maintaining a buffer between $\sum_{j \in J_i} a_{ij}x_j$ and b_i at optimality.

2.2 Ben-Tal and Nemirovski's ellipsoidal robust formulation

Ben-Tal and Nemirovski has made significant contributions to the field of robust optimization over the years. Of utmost significance is their approach to handle the over-conservative nature of Soyster's formulation by considering an ellipsoidal uncertainty set instead of a box uncertainty set as in Soyster's methodology. They proposed the following robust formulation:

$$
\begin{array}{lll}
\text{maximise} & c'x & \\
\text{subject to} & \sum_{j} a_{ij}x_j + \sum_{j \in J_i} \hat{a}_{ij}y_j + \beta_i(x) \leq b_i & \forall i \\
& \beta_i(x) = \Omega_i \sqrt{\sum_{j \in J_i} \hat{a}_{ij}^2 z_{ij}^2} & \forall i \\
& -y_{ij} \leq x_j - z_{ij} \leq y_{ij} & \forall i, j \in J_i \\
& l \leq x \leq u & \\
& y \geq 0. &
\end{array}
$$

where Ω is the "safety parameter".

This results in a formulation that has the property of being less conservative by excluding the possibility of all elements, a_{ij}, reaching the limits of their uncertainty intervals at the same time. The formulation however, has the undesirable property of being nonlinear and therefore more computational expensive to implement.

2.3 Bertsimas and Sim's adjustable robust formulation

Bertsimas and Sim proposed an approach that was not only linear and tractable but also allowed the user to control the degree of conservationism by varying the uncertainty budget, Γ. This uncertainty budget is defined row-wise as Γ_i and is a value that lies between 0 and cardinality of J_i while not necessarily being integer.

Bertsimas and Sim have shown that their formulation offers full protection against $\lfloor\Gamma_i\rfloor$ number of changes deterministically and a high probabilistic protection against further $(\Gamma - \lfloor\Gamma_i\rfloor)\hat{a}_{it}$ changes. They first consider the intermediate nonlinear formulation:

$$
\begin{aligned}
\text{maximise} \quad & c'x && && (2.1)\\
\text{subject to} \quad & \sum_j a_{ij}x_j + \max_{\zeta}\{\sum_{j\in S_i} \hat{a}_{ij}y_j + (\Gamma - \lfloor\Gamma_i\rfloor)\hat{a}_{it}y_{it}\} \leq b_i && \forall i && (2.2)\\
& -y_j \leq x_j \leq y_j && \forall j && (2.3)\\
& l \leq x \leq u && && (2.4)\\
& y \geq 0. && && (2.5)
\end{aligned}
$$

where ζ in Equation 2.2 is the set $\{S_i \cup \{t_i\} | S_i \subseteq J_i, |S_i| = \lfloor\Gamma_i\rfloor, t \in J_i \setminus S_i\}$.
This maximization term, also known as the protection function $\beta_i(x_i^*, \Gamma_i)$, can be reformulated as a linear programming model and its dual found to modify the above formulation so that it becomes linear.

The result is a linear robust formulation:

$$
\begin{aligned}
\text{maximise} \quad & c'x \\
\text{subject to} \quad & \sum_j a_{ij}x_j + z_i\Gamma_i + \sum_{j \in J_i} p_{ij} \leq b_i & \forall i \\
& z_i + p_{ij} \geq \hat{a}_{ij}y_j & \forall i, j \in J_i \\
& -y_j \leq x_j \leq y_j & \forall j \\
& l_j \leq x_j \leq u_j & \forall j \\
& p_{ij} \geq 0 & \forall i, j \in J_i \\
& y_j \geq 0 & \forall j \\
& z_j \geq 0 & \forall i.
\end{aligned}
$$

This formulation has the further benefits of being able to control the level of robustness for each constraint using Γ_i and being naturally applicable to discrete optimization problems. This formulation fulfils all the requirements that are being considered and was deemed to be the most suitable formulation to be applied to the problem which is presented in the next chapter.

Chapter 3

Model Framework and Computational Treatment

We describe in this section the forest planning model and the procedure of applying the robust formulation to obtain a robust counterpart. The original basic model used here is presented by Carrasco and Vera [8] but similar models have been developed by many other academia and industry modellers. The main decision of the basic model is to determine how much to harvest from each forest stand and where to transport it, given that there is a demand for a combination of log types at different production or export points (pulp plants, saw mills and ports for export). The main objective is to reduce the total cost which comprises of operational costs, transportational cost, assignment costs and a penalty cost for producing in excess. Extensions of this model deals with decisions related to uncertainty management and scheduling of machinery. The computational tools used to solve this problem is also discussed at the end of this chapter.

3.1 Sets, Variables and Parameters

The following sets, variables and parameters were used:

SETS

T: Time periods;

RD: All real demand destination points;

DD: Type of demand destinations {pulp, sawmill or export};

SD: Specific destinations for each type of demand destination in 'DD';

LT: All log types;

LD: Log type required by type of demand destination in 'DD';

AM: All harvesting machines;

MT: Type of machinery {Skidder, Harvester, Tower};

ME: Harvesting equipment of type 'MT';

FS: All forest stands;

SMT: Stands harvestable by machines in 'MT';

DECISION VARIABLES

X_{isjt}: Amount of surface to be harvested in stand i to demand destination j of type s in period t;

Z_{imt}: binary decision variable to determine if machiner m is active in stand i in period t;

W_{ipmt}: binary decision variable to determine if machine m will move from stand i to stand p after period t;

AUXILIAR VARIABLES

SC: Sum of penalty, operational, transportation and assignment costs;

PC: Penalty cost of not meeting demand;

OC: Operational cost of harvesting;

TC: Transportation cost of transporting logged trees to demand destinations;

AC: Assignment cost of assigning harvesting machinery to each forest stands;

PARAMETERS

MC: Machinery capacity;

$disc_t$: Discount rate in period t;

PL_k: Price of each log of type k ;

SA_i: Surface available in stand i;

HY_{ik}: Harvest yield of logs of type k per hectare of surface in stand i;

hc_{im}: Unit harvesting cost in stand i with machine m;

tc_{id}: Unit transportation cost of logs from stand i to destination d;

ac_{ijq}: Unit assignment cost of routing from stand i to stand j with machine type q;

Dd_{sjkt}: Demand of log type k with destination j of destination type s at time t;

wo: Weight on the harvesting cost;

wt: Weight on the transportation cost;

wa: Weight on the assignment cost;

P_Ex: Penalty for excess in the demand;

Γ: Uncertainty budget;

α: Percentage deviation for each yield parameter;

INDEXING CONVENTION

For clarity in presenting the model, the rest of this thesis will adopt the following convention for indexing. All subscripts are indexed to their respective sets as defined:

$$
\begin{aligned}
t &\in T \\
s &\in DD \\
j(s) &\in SD \\
k(s) &\in LD \\
i &\in FS \\
q &\in MT \\
i(q) &\in SMT \\
p(q) &\in SMT \\
m(q) &\in ME
\end{aligned}
$$

3.2 Harvesting model description

The basic model is:

minimise SC

subject to:

$$PC + OC + TC + AC \leq SC \tag{3.1}$$

$$PC = P\_Ex \sum_{t} disc_t * \Big(\sum_{s} \sum_{j(s)} \sum_{k(s)} PL_k (\sum_{i} HY_{ik} * X_{isjt} - Dd_{sjkt}) \Big) \tag{3.2}$$

$$TC = \sum_{t} disc_t * (\sum_{i} \sum_{s} \sum_{j(s)} \sum_{k(s)} wc * tc_{ij} * HY_{ik} * X_{isjt}) \tag{3.3}$$

$$OC = \sum_{t} disc_t * (\sum_{q} \sum_{i(q)} \sum_{s} \sum_{j(s)} \sum_{m(q)} wo * hc_{im} * X_{isjt}) \tag{3.4}$$

$$AC = \sum_{t} disc_t * (\sum_{i} \sum_{j} \sum_{q} \sum_{m(q)} wa * ac_{ijq} * W_{ijmt}) \tag{3.5}$$

$$\sum_{s} \sum_{j(s)} \sum_{t} X_{isjt} \leq SA_i \qquad \forall i \tag{3.6}$$

$$\sum_{i} HY_{ik} * X_{isjt} \geq Dd_{sjkt} \qquad \forall s, j(s), k(s), t. \tag{3.7}$$

Constraints (3.1) to (3.5) is the breakdown of the objective function into the respective components for analysis in determining the relative importance of the costs. Constraint (3.6) states that the total amount harvested over the entire time period should be less than the total available resources. Constraint (3.7) states that the amount harvested must be able to meet the demand.

In the forest industry, the parameters that are most susceptible to uncertainty are the harvest yield and the amount of demand. As only short term decisions are analyzed in this thesis, the demand can be assumed to be certain based on the fact that contracts for orders are enforced in the short term. However, in considering models spanning a longer timescale, demand uncertainty effects have to be considered. The harvest yield parameter appears in 3 sets of constraints; 2 of which are the objective components written as constraints (3.2) and (3.3) and 1 of which is the demand constraint (3.7). The uncertainty can also come from other sources besides differing growth such as pests, disease, fires or even heavy rain.

Conflicts existing between the worst case scenarios of the objectives and the demand constraints have been previously pointed out and studied. The worst case for the demand constraints is when there is a decrease in the harvest yield parameters but the worst case scenario for the objective function constraint is when there is an increase in the harvest yield parameters. This leads to a solution which overcompensates for this conflicting interest by taking an overly conservative solution.

The study in [16] shows the differences in objective value when considering the inclusion of either only the demand constraint or both the demand and objective function constraints. Although there is a difference in the objective value obtained in both approaches, the mean difference is approximately 10% over the range studied in this work. Based on the argument that this difference is not too large, slightly over conservative and to eliminate the computational effort required to model the 2 conflicting constraints accurately, this research will only consider the robust counterpart of the demand constraints.

Applying the robust formulation on constraint (3.7) , the variables Y_{isjt}, q_{sjkt}, and p_{isjkt} are first created and a common Γ value $\in$ [0,12] for every row is set instead of setting a Γ_i for every row. The robust reformulation for this constraint based on Bertsimas and Sim's robust formulations thus becomes:

$$\sum_i HY_{ik} * X_{isjt} - q_{sjkt} * \Gamma + \sum_i p_{isjkt} \leq Dd_{sjkt} \qquad \forall s, j(s), k(s), t \tag{3.8}$$

$$q_{sjkt} + p_{isjkt} \geq \alpha * HY_{ik} * Y_{isjt} \qquad \forall i, s, j(s), k(s), t \tag{3.9}$$

$$0 \leq X_{isjt} \leq Y_{isjt} \qquad \forall i, s, j(s), t \tag{3.10}$$

$$Y_{isjt} \geq 0 \qquad \forall i, s, j(s), t \tag{3.11}$$

$$p_{isjkt} \geq 0 \qquad \forall i, s, j(s), k(s), t \tag{3.12}$$

$$q_{sjkt} \geq 0 \qquad \forall s, j(s), k(s), t. \tag{3.13}$$

where α in constraint (3.9) is a non-negative parameter used to represent the uncertainty distribution in terms of the original yield parameter and the convention for indexing follows from previously. Constraints (3.8) to (3.10) are constraints obtained by writing the robust counterpart of the constraint (3.7) and therefore replace (3.7) in the final model. Constraints (3.11) to (3.13) are the non-negativity constraints on the dual variables.

3.3 Routing Extensions

The variables and constraints that were developed to handle the routing considerations are discussed in this section. The problem at hand is one that is analogous to a multi-period plant location problem with a fixed charge network flow problem with both being NP-hard problems that are difficult to solve. Much effort has been expended in solving these problems to optimality as detailed in the Section 1.2. If the machines are allowed to move unrestricted to any stand as an intermediate point to the next stand which they harvest, this translates into solving a Steiner tree problem which can have many solutions for an initial set of vertices. For easy reference, this formulation is named Formulation A. The Steiner problem is NP-complete and also hard to approximate. Approximation algorithms have been investigated and best upper bounds at 95/94 of the optimal value have been found but this best approximation algorithm is NP-hard as well. [9].

An engineered approach is therefore needed to obtain solutions which are at least obtainable within a reasonable time. Many different variations of formulations were created and tested but eventually the key insight came from the earlier mentioned fact that each type of machinery is only suitable for harvesting certain terrains and therefore certain stands. Hence, by letting the machines move only to stands which they can harvest directly and not allowing any intermediate points, we are able to avoid solving the Steiner tree problem. Although this may seem more restrictive, it makes some operational sense as leaving machinery in stands which they cannot harvest also incurs opportunity cost which have not been modelled into the original model as well as incurs extra costs from splitting the shifting of machinery over several periods. This formulation is named Formulation B and the comparison of the computational effort in solving these two approaches is shown in Table 3.1.

Table 3.1: Comparison of routing formulations

Case	No. of constraints	Time(s)	No. of nodes
Nominal	158	0.0468003	-
Formulation A	15570	27353.68	514378
Formulation B	3318	295.294	94403

For a given data test instance, the nominal problem without routing solves almost instantly whereas Formulation A does not solve to optimality even after approximately 7.6 hours with a relative MIP gap of 13.4% at user termination and has a large number of brand and bound nodes. Formulation B on the other hand does well and solved to optimality in a reasonable time of 295.3 seconds. The formulation for the second approach is thus used as a means of obtaining reasonable routing decisions in a reasonable time and is detailed here.

The feasible region is defined by the constraints (3.14) to (3.19) as well as the binary variables Z_{imt} and W_{ijmt}. Binary variable Z indicates which stands are harvested by machinery during each period and variable W indicates the movements of the machinery with their corresponding origin and destination pair. Mathematically, this takes the form of

$$Z_{imt} = \begin{cases} 1 & \text{if stand i is harvested by machine m in time period t,} \\ 0 & \text{otherwise.} \end{cases}$$

$$W_{ipmt} = \begin{cases} 1 & \text{if machine m moves from stand i to stand p after time period t,} \\ 0 & \text{otherwise.} \end{cases}$$

The routing constraints that are added to extend the model to consider machinery scheduling issues are:

$$\sum_{s}\sum_{j(s)} X_{isjt} \leq MC * \sum_{m(q)} Z_{imt} \qquad \forall q, i(q), t \tag{3.14}$$

$$\sum_{m} Z_{imt} \leq 1 \qquad \forall i, t \tag{3.15}$$

$$\sum_{i(q)} Z_{imt} \leq 1 \qquad \forall q, m(q), t \tag{3.16}$$

$$Z_{imt} \leq \sum_{p(q)} Z_{pm,t+1} \qquad \forall q, i(q), p(q), m(q), t \tag{3.17}$$

$$Z_{im,t+1} \leq \sum_{p(q)} Z_{pmt} \qquad \forall q, i(q), p(q), m(q), t \tag{3.18}$$

$$Z_{imt} + Z_{pm,t+1} \leq W_{ipmt} + 1 \qquad \forall q, i(q), p(q), m(q), t. \tag{3.19}$$

Constraint (3.14) states that the maximum area harvested at each stand i harvestable by machine m of class q is limited by machine capacity, MC, per period. Constraint (3.15) prevents the assignment of more than 1 crew to stand i at period t. Constraint (3.16) states that only 1 stand harvestable by the type of machine in a particular machine class q can have machine m of the same machine class q at each time period. For constraints (3.17) to (3.19), the index for time period, t, runs from the initial period to the second last period so as not to exceed the time boundaries due to the presence of the expression t+1 in the constraints. Together, constraints (3.17) and (3.18) define the relations that if a machine is being used in any one period, then there should be a previous path and a following path which will result in a final path from the initial time period to the final time period. Constraint (3.19) forces the variable W to take on a value of 1 if both Z_{imt} and $Z_{im,t+1}$ take on a value of 1.

Due to the presence of W in the minimizing objective function, W will take a value of 0 otherwise. Using this insight in practice, the binary variable W is relaxed to be $0 \leq W \leq 1$ but it will still only take values of 0 or 1. This results in a substantial decrease in the number of binary variables in the model that aids in the use of a branch and bound approach to solve the problem. Together, all these constraints constitute the required requirements for obtaining the optimal paths for the harvesting machinery in the multiperiod harvesting problem and are added to the final model.

3.4 Computational Treatment

The models in this work were modeled using AMPL and solved with CPLEX 10.1. The hardware used was an IBM Thinkpad with Intel Duo Core CPU 2.0 Ghz and 2.0 Gb of RAM, under the Windows Vista operating system.

AMPL is an advanced modelling language for large-scale optimization and mathematical programming problems. AMPL allows for an easy translation from familiar mathematical and algebraic notation to programming code and allows for a high level of user control through their interactive command environment. The author has found very useful the ease in defining multidimensional data and nested sets and the flexibility in scripting and solution reporting. In general, it allows for easy definitions of a large myriad of complex optimization programs in concise forms and is able to link up to a wide variety of powerful solvers including the CPLEX optimizer.

CPLEX 10.1 is a state-of-the-art solver which utilises algorithms that implements the dual simplex, primal simplex, network simplex algorithms and barrier methods. Of much relevance and significance to this work is their mixed integer optimizer which includes sophisticated pre-processing algorithms and many strategies that utilizes cutting planes and heuristics for finding the optimal integer values. All these strategies and processing can be readily customized by the user. The author has found the ability to customize search preferences in the branch and bound approach and warm-starting mixed integer programs useful in finding the optimal values of NP-hard combinatorial problems.

Chapter 4

Case Motivation and Computational Results

4.1 Case Motivation and Framework

Several instances were created in order to evaluate the behavior of the robust solution to the model with routing constraints. The motivating questions to the creation of these instances are:

1. How does the solution to the robust problem change with respect to changes in the uncertainty budget and expected variability?

2. Does the structure of the solution change with the addition of robust and/or routing considerations? If so, how?

3. How does the robust or routing solutions behave when given an *a posteriori* uncertain model?

The first 2 questions are addressed in this chapter and the last question discussed in chapter 5. The robust approach is evaluated in a variety of cases which depend on industrial or slightly modified industrial data. The problem instances are significantly smaller than industry standards in order to obtain solutions in a reasonable time but they are easily scalable to the required industrial dimensions. Our problem consists of

12 parcels of forest, 6 destinations, 8 types of tree species, and 8 harvesting machines and spans a total of 12 harvesting periods.

4 cases were implemented in order to answer the questions stated above:

Case 1: This is the nominal robust case without routing constraints using data originating from a Chile forest harvesting firm. The variabilities are set at 5%, 10% and 20% with respect to the nominal values to determine the changes in the objective function and structure of the solution as the uncertainty budget is increased sequentially from 0 to the maximum of 12. As the variability increases, it is expected that the model will take a more conservative solution to retain its robustness with respect to changes in the harvest yield.

Case 2: This is a modified version of Case 1 whereby routing considerations are added and the solution structure and harvesting decisions are studied to determine the resulting changes compared to Case 1. It is expected that with routing decisions and assignment costs added to the model, the optimal objective value will naturally increase. Another point of interest is whether the presence of routing constraints will affect the initial harvesting decisions when variability and Γ are the same.

Case 3: This is a modified version of Case 1 whereby two stands are made equal in all aspects but will be subjected to different variability. This is a simple exercise to exemplify that the model will make decisions in a rational manner and chose to harvest the stand with lesser variability so as to reduce the risk due to uncertainty, ceteris paribus.

Case 4: This is a modified version of Case 3 where routing decisions are added. We compare the solution for this case with the solution from Case 3 when the variability is set to be the same. This is another simple example that clearly shows the differences in solution structure due to the presence of routing considerations, ceteris paribus.

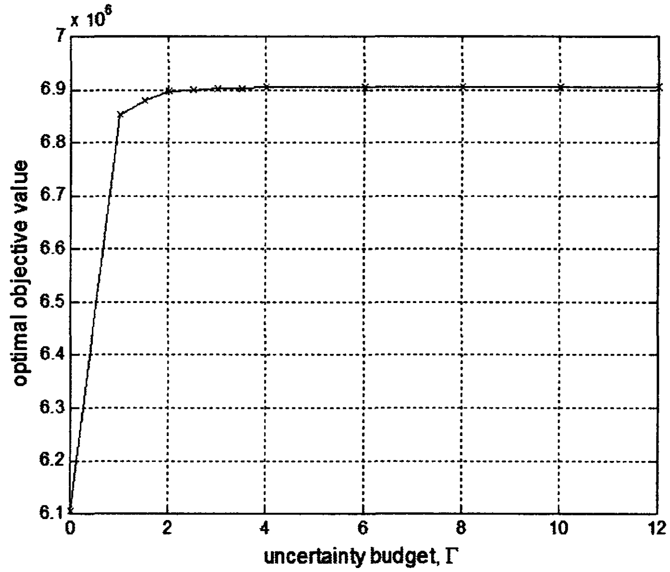

Figure 4-1: Objective values of robust model at 5% variability as a function of Γ

4.2 Computational Results

4.2.1 Case 1: Basic Robust Case

Case 1 uses the original industrial data to obtain the optimal multi-period harvesting decisions when protected against uncertainty. Variabilities of 5%, 10% and 20% are introduced into the harvest yield and the objective values are tabulated and plotted. From Figure 4-1 which shows the optimal objective values at 5% variability and from Figure 4-2, several phenomenons can be observed. Firstly, it is noted that it only took a very small increase in the amount of protection for the objective value to peak and plateau. The optimal objective value remains fairly constant for an uncertainty budget value of 4 and above. This implies that sufficient protection against the worst case scenario is provided with a Γ value of 4 and there is no need to budget for a higher level of protection for this particular set of data. Optimal objective plots for other cases of variability at 10% [B-2] and 20% [B-3] exhibit the same behavior and can be found in Appendix B.

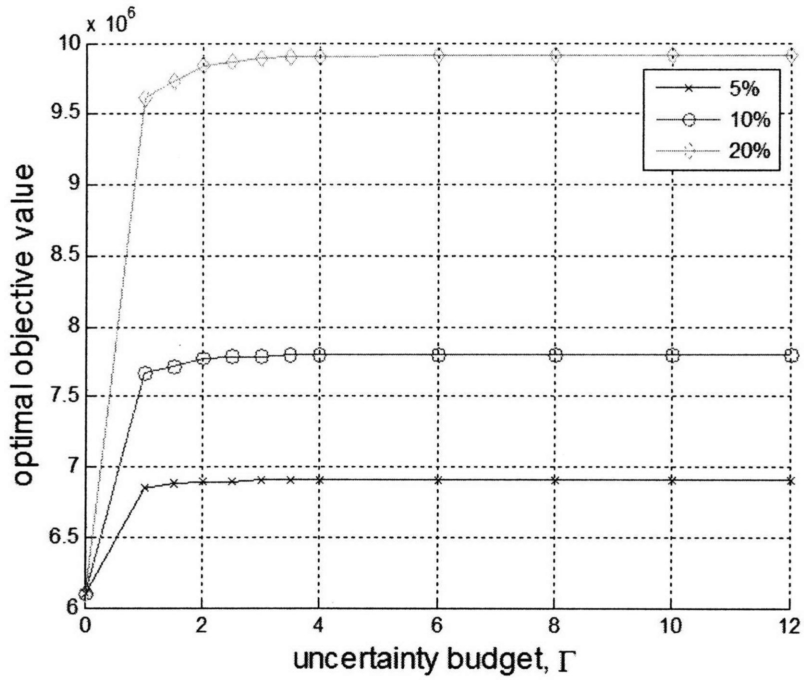

Figure 4-2: Comparison of objective values of the robust model at different variability as a function of Γ

The results indicating a worst case scenario being attained at a Γ value of 4 is reasonable, given that the solutions shows that the top 4 producing forest stands produce approximately 85% of the total required demand. A sample breakdown of the production by stands for all 3 formulations are shown in Table 4.1.

From Figure 4-2, it can be seen that as the variability increases, the objective value increases as well, reflecting a cost increase due to an increase in the level of harvesting. As can be seen in Table 4.3, in transiting from the nominal model to the robust model, the model decides to harvest more at less risky locations and less in other locations in order to protect against uncertainty. Overall, the model harvests more in total to protect against uncertainty and meet the total required demand. Case 3 does a simple exemplification of this effect by comparing 2 stands modified to be similar but are subjected to different levels of uncertainty.

Table 4.1: Breakdown of output by stands for $\Gamma = 2$ and 5% variability

Stand	Nominal%	Robust%	Routing%
1	17.2	16.5	18.7
2	0	0	0
3	16.2	16.2	16.5
4	3.1	3.1	4.4
5	2.3	2.3	3.09
6	2.7	2.7	2.8
7	9.2	9.7	8.4
8	40.9	40.8	40.4
9	0.8	0.9	0
10	3.0	3.2	2.9
11	1.9	1.9	0.0
12	2.7	2.6	3.0
Total	100	100	100

Table 4.2: Breakdown of costs averaged over all Γ for 5% variability

Stand	Robust	Routing	Percentage increase
Penalty	481558.4	589882.9	22.5
Operational	1380803.1	1379738.3	-0.1
Transportation	5033166.8	5418603.8	7.7
Assignment	-	42346.6	-
Total	6895528.2	7434666.9	7.8

4.2.2 Case 2: Robust Case with Routing Considerations

With the addition of the routing constraints, the objective value increases from the previous case by 7.5% on average over all the scenarios considered as seen in Figure 4-3 for the case with 5% variability.

From Table 4.2, the breakdown of the total costs can be observed for 5% variability in harvest yield. At first sight, the increase due to assignment cost is not large and does not justify the modeling and computational efforts to include the routing constraints. However, the influence of the routing constraints leads to increases in all components of cost. For example, in the case shown, the total costs actually in-

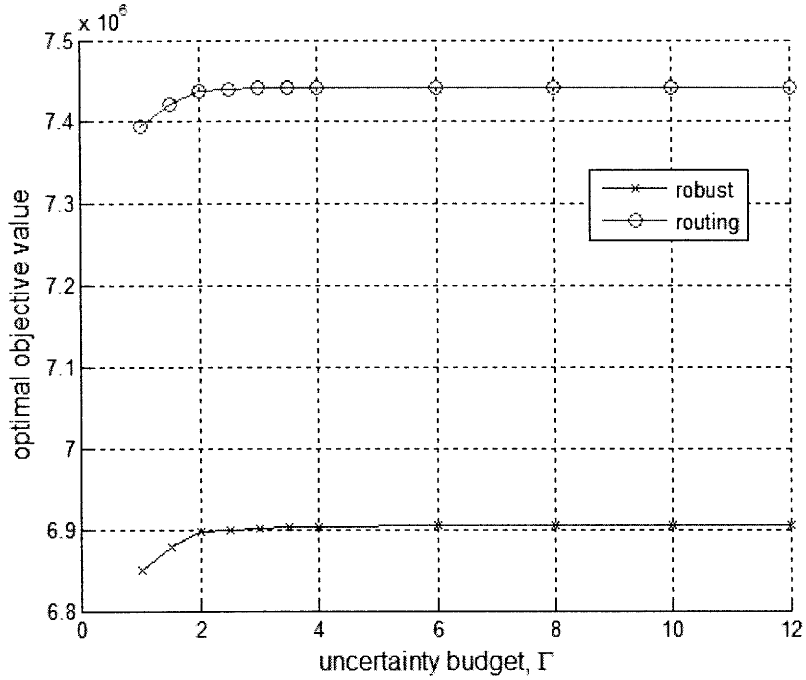

Figure 4-3: Objective values of robust model at 5% variability as a function of Γ

creased by 7.8% which is mainly due to the increase in transportation costs. This can be explained by realising that due to the fact that harvesting is being restricted to areas routed by the machines, certain forest stands with cheap transportation to final demand destinations could not be harvested. Data for other variabilities show the same breakdown and is shown in the Appendix A as Tables A.5, A.6 and A.7. This indirect cost increase due to the inclusion of routing constraints is a large motivation for modelling the routing constraints to get an accurate optimal decision and is substantial considering the absolute cost in hundreds of millions of this industry.

Table 4.3: Percentage distribution of harvest production for $\Gamma = 2$ and 5% variability at time period 5

Stand	Nominal	Robust	Routing
1	14.4	7.02	31.49
2	0	0	0
3	0	0	0
4	0	0	14.30
5	0	0.15	0
6	0.16	0.15	1.68
7	8.14	14.49	0
8	53.74	52.14	51.28
9	0	1.55	0
10	1.36	3.86	0
11	21.14	20.64	0
12	1.06	0	1.25
Total	100	100	100

An interesting question is whether the addition of the routing constraints will result in any change of the solution structure when compared to the previous case. It is noted from the data as shown in Table 4.3 that the solution structure does change. After adding routing considerations, the model harvests an even larger amount but at a smaller number of stands. Case 4 examines a modified model that helps to explain a plausible reason for this change in the solution structure.

Table 4.4: Sample data for a comparison of similar stands subjected to different variability

Γ	0.0	1.0	1.5	2.0	2.5	3.0	3.5	4.0
A, Stand 2	0	0	0	0	0	0	0	0
A, Stand 8	100.6	101.3	103.5	104.3	105.5	105.9	105.9	105.9
B, Stand 2	0.0	41.3	40.2	30.6	17.5	0.0	0.0	0.0
B, Stand 8	100.6	38.5	41.1	57.4	68.4	105.9	105.9	105.9

4.2.3 Case 3: Robust Case with 2 identical stands

In order to test the behavior of the robust approach when the yield coefficients are subjected to different uncertainty, a case with two almost exactly identical forest stands were created with the exception that their yield variability vary. Scenario A was created whereby stand 2 and stand 8 were chosen to be identical and their harvest yield was subjected to 5% variability. Given the case when their yield variability is exactly the same, the model does not differentiate between the two stands and divides the harvesting amount equally between the two or randomly select either one to harvest the full amount. In Table 4.4, it can be seen that the solver chose to harvest from stand 8 for this case. For scenario B, Stand 2 was then subjected to a higher degree of variability. In the case when one stand has a higher variability than the other competing stand, the model will always choose the stand with less variability to harvest as long as the demand is being able to be supported by that single stand. In the event that demand outstrips the supply from a single stand, the model harvests from the stand with least variability first before harvesting from the other stand. As an illustration, the model was allocated different uncertainty budgets ranging from 0 to 12 and the result is shown in Table 4.4.

The model shows interesting results when the allocated uncertainty budget was varied. The model decides to harvest more at stand 8 which has lower variability as Γ is increased. This is in agreement with what one will expect when the model tries to avoid risk from uncertainty.

Table 4.5: Simple Illustration of change in harvest decision due to routing constraints

Nominal	Period 1	Period 2	Period 3	Period 4	Period 5
R2	50.3134	70.6214	70.6214	70.6214	35.0927
R8	50.3133	70.6214	70.6214	70.6214	35.0927
Robust					
R2	52.9614	74.338	74.338	74.338	36.9397
R8	52.9614	74.338	74.338	74.338	36.9397
Routing					
R2	111.807	156.937	156.937	156.937	77.9839
R8	0	0	0	0	0

4.2.4 Case 4: Robust Case with Routing considerations and 2 identical stands

A similar idea is extended to test the effect of routing considerations on two identical stands. In this case, the yield variability of both stands 2 and 8 are set to be the same which resulted in an equal distribution of forest harvesting from both stands initially and can be seen from the data exhibited in Table 4.5. Stands 2 and 8 were only harvested from periods 1 to 5 and the rest of the periods were excluded from the table. It is noted that with the addition of routing constraints, the model will chose to harvest at just one of the locations. This can be explained by understanding from a cost approach. By only harvesting from only one stand instead of two, the model avoids costly assignment costs for 2 machines and instead assigns the harvesting to just 1 machine. This is also one of the reasons for the change in the structure of the unmodified industrial data in Case 2. The model tries to reduce the number of machines used to reduce costs and the result is a smaller amount of machines used for harvesting larger aggregated amounts of logs in a smaller number of stands.

This actually leads to a better overall solution to the harvesting problem from an operational viewpoint. When running an optimization model, it might be numerically optimal to harvest in small amounts from various stands but this might be difficult to implement in reality as other costs for crew and overheads not originally included but have to be taken into account in practice might result in such small amounts becoming uneconomical to harvest. In other words, the addition of the routing constraints will result in an aggregation of harvesting to a smaller number of stands, reducing the number of cases whereby several small amounts are harvested from many stands that might be hard to fulfill or may lead to higher costs operationally.

Chapter 5

Monte Carlo Feasibility Study

5.1 Description

In order to justify the use of a robust methodology to industrial use, and to determine the probability of a computed solution being feasible in practice, a Monte Carlo study of feasibility was implemented. The study aims to provide evidence for the usability of the solution in making decisions in actual practice and encourage the use of a robust methodology for forest harvesting planning among practitioners. To implement a Monte Carlo study of feasibility, 2000 different scenarios for the yield coefficients were generated whereby each parameter of the harvest yield was perturbed by sampling from a uniform distribution within the uncertainty interval. The solutions obtained using the basic robust and routing formulations were then applied to these 2000 scenarios to determine if they still satisfied the required demand. A record of the percentage of feasible scenarios and the actual amount of infeasibility for each level of the uncertainty budget was made. The average amount of infeasibility actually corresponds to the average amount of unmet demand whenever infeasible and this amount is obtained by averaging the total amount of infeasibility over the number of cases which are infeasible. Lastly, the average objective value for the basic robust and routing formulations when applied to the 2000 scenarios was calculated and compared with the original to determine the difference.

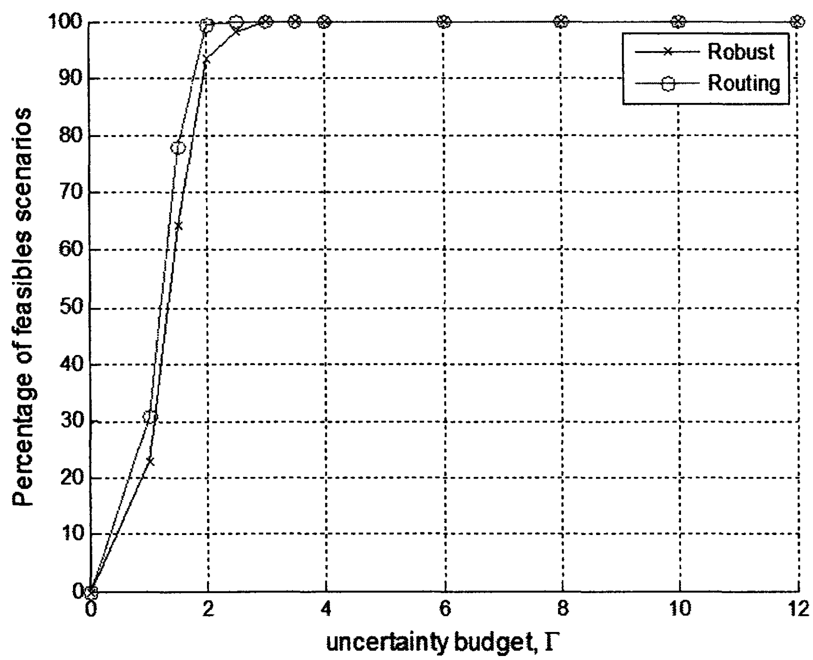

Figure 5-1: Percentage of feasibile senarios at 5% variability as a function of Γ

5.2 Results

It is immediately noted that in all cases of different variability, the percentage of feasible cases increases as we increase the uncertainty budget. The can be seen from Figure 5-1 and plots for other variabilities B-7, B-8, and B-9 in Appendix B. This confirms the previous analysis that the worst case is protected by an uncertainty budget with Γ equal to 4 as the percentage of feasible scenarios reach 100% or close to 100% when as Γ is increased to 4. It is also interesting to note that the percentages of feasible cases for the routing models are almost always slightly greater than the percentages of feasible cases for the basic robust models. This can be explained by understanding that due to the presence of the routing constraints, the solution is forced to take on a more rigid structure which includes aggregation of harvesting quantities. This aggregation of harvesting decisions to certain stands will result in less variability by limiting the number of harvested stands that are affected by variability and can be thought to have the same effect as risk pooling.

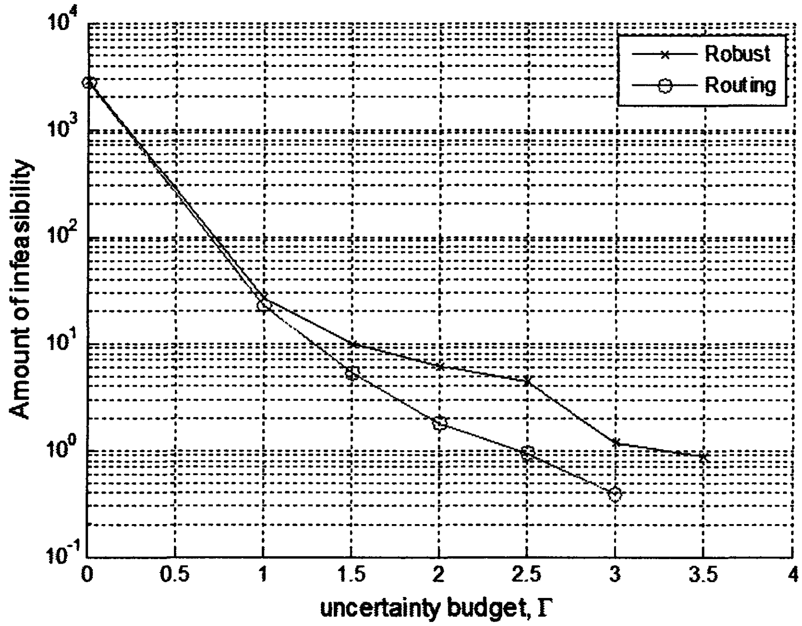

Figure 5-2: Average infeasibility of senarios at 5% variability as a function of Γ

One seeming anomaly is that the nominal solution has zero or a low percent feasibility for all the simulations created. One would expect the nominal solution to be at least feasible for some scenarios. However, this can be explained by understanding that the parameters in the harvest yield are allowed to vary individually. The probability of all constraints moving in the same direction so as to render the original nominal solution feasible is close to zero due to the large number of constraints and hence strengthens the need for a robust formulation.

From the graphs of the average infeasibilities Figures 5-2, B-11 and B-12, several insights can be drawn. Most notable is the fact that the amount of infeasibility or unmet demand is very high when one does not consider the use of a robust solution at all. This unmet demand is rapidly reduced by 2 orders even when only a slight protection is enabled using Γ equal to 1. This strongly encourages the use of robust optimization even at a very low level in order to reduce the cost associated with not meeting demand which might include failure to meet contractual agreements and lost of goodwill besides the lost opportunity cost of making more profit. The other insight is that once protection against uncertainty takes place, the unmet demand is actually only a small proportion of total demand. This leads to the concept of a "weak" infeasibility as opposed to a "strong" infeasibility.

A "strong" infeasibility condition is defined in this paper as the condition that a scenario is deemed infeasible as long as it does not meet demand whereas a "weak" infeasibility condition is when the unmet demand is not excessively large and can possibly be handled in practice using safety inventory or purchase from external suppliers in order to meet the contractual obligations. Hence, an industrial practitioner can decide to protect up to a certain amount based according to a mix of safety inventory standards and probability. For example, if an industrial practitioner feels that it is reasonable to meet demand 95% of the time, then he or she can make use of the data to determine the suitable Γ value that will result in a solution that leads to a 95% chance of feasibility and obtain the harvesting decisions using that budget of uncertainty. Alternatively, he or she can make use of the average infeasibility data to compare with the available safety inventory to chose a budget for uncertainty that is used to obtain harvesting decisions.

Another interesting question that might arise would be to determine the differences in the amount of feasibility if more information on the distribution was known. For example, the user might want to set the level of protection against variability within each parameter to be a certain percentage of a normal distribution. Working with a 95% confidence interval that the sample will be within the protection level of 5% of the nominal data, the following calculations results in the required formula for variance of each parameter. From the normal distribution tables, a 95% confidence interval corresponds to 1.96σ where σ is the standard deviation of the parameter.

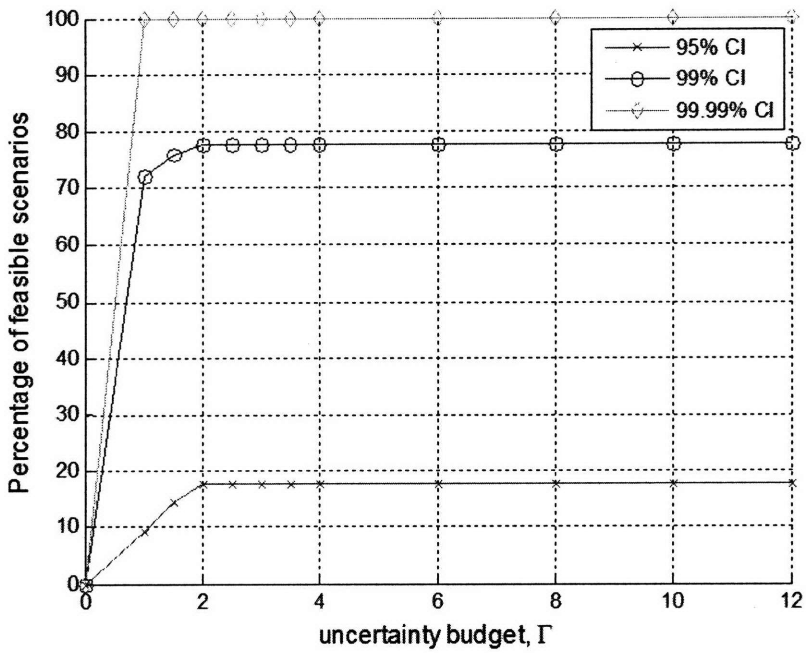

Figure 5-3: Comparison of feasibility for different confidence intervals with respect to Gamma

Setting this confidence interval to be equal to the level of protection that is given:

$$
\begin{aligned}
1.96\sigma &= 0.05 * HY_{ik} \\
\sigma &= \frac{0.05 * HY_{ik}}{1.96} \\
\text{variance, } \sigma^2 &= \left(\frac{0.05 * HY_{ik}}{1.96}\right)^2
\end{aligned}
$$

Sampling from the normal distribution, one arrives at Figure 5-3 where it is observed that the graphs rises rapidly but reaches a plateau for each level of confidence interval used. This is actually within expectations as the normal distribution is distributed closer to the center as compared to a uniform distribution and therefore there will be a higher level of feasibility for the whole range of protection at 99.99% confidence interval. However for the other 2 graphs, due to the fact that protection is limited to 95% or 99% of the distribution, the probability of all sampled parameters being able to fulfil all the contraints is greatly reduced which leads to a high chance of infeasibility. However, it is also noted that the amounts of infeasibility in these cases are very low which can be attributed to the tails of the normal distribution. Given that most operational processes have an underlying uncertainty distribution which is empirically close to a normal distribution, inferences can be drawn from the 6-Sigma methodology to this test to understand that uncertainty containment have repercussive effects and an adequate amount of variability protection must be used in conjuction with protection against number of parameter changes.

Chapter 6

Conclusion

6.1 Summary

In this thesis, the key points established from the series of industrial and experimental cases are that the use of a robust approach for planning forest operations is important and essential to safeguard against uncertainty induced by the biological nature of the forest industry as well as other external factors. The extension to include routing considerations is vital for an accurate modeling of the actual forest harvesting practice.

Given a solution that is not protected against uncertainty, the possibility of it being infeasible in practice is very high and the resulting shortage is usually counterbalanced by harvesting more than required based on experience from a planner or by using safety inventory. However, the decisions made based on this approach are usually dependent on the planners' judgment and experience. On the other hand, the use of a robust formulation helps in protection against unmet demand deterministically. Furthermore, by using the data obtained from a Monte Carlo study of feasibility, a practitioner can judge the amount of uncertainty protection to use for a given probability of supply shortage scenarios or available safety inventory and then determine the required harvesting decisions using a robust routing model.

The addition of routing considerations results in a larger problem that is NP-hard and takes a significantly longer time to solve to optimality but is shown to be important due to the substantial costs that are included when considering the use of the capital intensive machinery and also due to its effects on the way forests are harvested. It was also noted that the consideration of routing constraints may lead to other operational benefits such as a reduction in harvest decisions that are uneconomical in practice due to their limited amounts. The inclusion of the routing constraints is therefore strongly recommended in order to obtain an enhanced model.

Even after some simplifications that reduce the complexity of the problem, the routing problem remains NP-hard and highly data dependent. Although computational time and effort for the industrial data set is reasonable, investigation into other data sets may require more time or effort in tuning the CPLEX options to yield solutions in reasonable time. This brings about the need for other approaches to evaluate the effects of a robust formulation on large scale routing problems and has brought forth interesting ideas for further in-depth research in this field.

6.2 Future research directions

Suggestions for future work include the use of a heuristic approach such as those discussed in section 1.2 to finding the routing decisions. The application of specialized heuristics to obtain solutions to routing problems have been successfully and extensively used for many years and a good starting reference is Simchi Levi's The Logic of Logicians [17]. The author's reason for not using a heuristic initially is due to the fact that a heuristic method will result in only close to optimal solutions generally and result in a more protected solution naturally. The solution hence undermines the effects of a robust formulation which is a major study of this research. However, as the data size increases, the "curse of dimensionality" manifests and a branch and bound approach as used by CPLEX will be unrealistic in practice. A study for large scale systems with this trade off in mind might be further pursued if required.

Appendix A

Tables

Table A.1: Objective values of robust model for various Γ and variability

Γ	5%	10%	20%
0	6102947.528	6102947.531	6102947.528
1	6850948.191	7663581.854	9609709.937
1.5	6878003.959	7718702.813	9728387.983
2	6896994.246	7770591.851	9845051.404
2.5	6899147.785	7781390.118	9869469.098
3	6901294.475	7786270.592	9887177.811
3.5	6902532.86	7790210.145	9897025.647
4	6903506.324	7792834.171	9904455.824
6	6904595.563	7795315.597	9911040.943
8	6904595.561	7795315.598	9912182.465
10	6904595.562	7795315.597	9912258.293
12	6904595.563	7795315.596	9912258.293

Table A.2: Objective values of routing model for various Γ and variability

Γ	5%	10%	20%
0	6615132.485	6615132.485	6615132.485
1	7395206.599	8236262.504	10275256.47
1.5	7420799.82	8292771.378	10396026.09
2	7437134.159	8340241.862	10509536.81
2.5	7439419.971	8354083.436	10558162.38
3	7441093.652	8358277.976	10595631.39
3.5	7441281.074	8358938.608	10612394.29
4	7441290.82	8359238.601	10622057.22
6	7441292.199	8359249.177	10622058.01
8	7441293.865	8359249.156	10622060.87
10	7441230.789	8359249.155	10622059.07
12	7441292.653	8359249.155	10622058.22

Table A.3: Average infeasibility of robust model for various Γ and variability

Γ	5%	10%	20%
0	-2845.871269	-5693.960968	-11391.44567
1	-26.843563	-51.751349	-96.576993
1.5	-9.820904	-19.217695	-36.377346
2	-6.137512	-14.749244	-33.639074
2.5	-4.514485	-9.16407	-18.800925
3	-1.188583	-4.530578	-8.818877
3.5	-0.888563	0	-0.512733
4	0	0	0
6	0	0	0
8	0	0	0
10	0	0	0
12	0	0	0

Table A.4: Average infeasibility of routing model for various Γ and variability

Γ	5%	10%	20%
0	-2776.640292	-5571.279827	-11201.75657
1	-22.221341	-42.945354	-81.513159
1.5	-5.368829	-17.585026	-31.809073
2	-1.80353	-11.010896	-18.459489
2.5	-0.940033	-2.013524	-7.102701
3	-0.388653	-0.657618	-4.381674
3.5	0	0	0
4	0	0	0
6	0	0	0
8	0	0	0
10	0	0	0
12	0	0	0

Table A.5: Breakdown of costs averaged over all Γ for 5% variability

Stand	Robust	Routing	Percentage increase
Penalty	481558.4	589882.9	22.5
Operational	1380803.1	1379738.3	-0.1
Transportation	5033166.8	5418603.8	7.7
Assignment	-	42346.6	-
Total	6895528.2	7434666.9	7.8

Table A.6: Breakdown of costs averaged over all Γ for 10% variability

Stand	Robust	Routing	Percentage increase
Penalty	1007022.8	1125594.7	11.8
Operational	1459806.3	1461419.2	0.1
Transportation	5304520.4	5706130.3	7.6
Assignment	-	46565.9	-
Total	7771349.5	8339710.1	7.3

Table A.7: Breakdown of costs averaged over all Γ for 20% variability

Stand	Robust	Routing	Percentage increase
Penalty	2254427.5	2400202.0	6.5
Operational	1637902.7	1648018.5	0.6
Transportation	5961217.0	6455175.3	8.3
Assignment	-	47268.0	-
Total	9853547.1	10550663.7	7.1

Appendix B

Figures

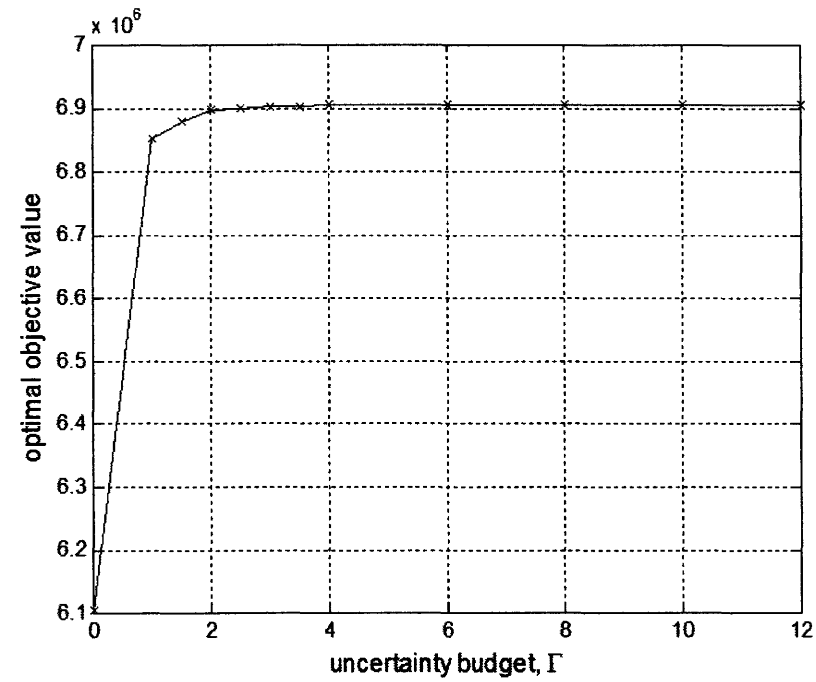

Figure B-1: Objective values of robust model at 5% variability as a function of Γ

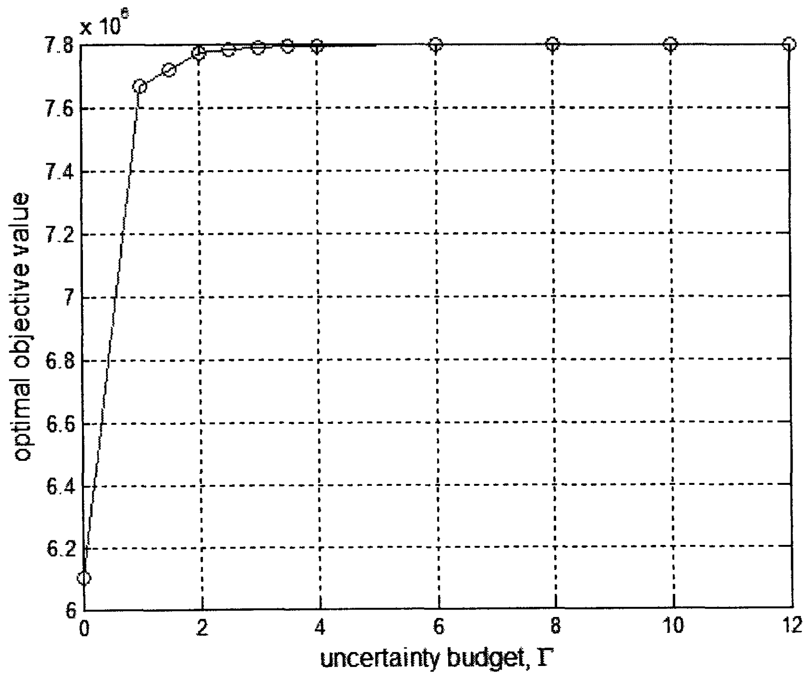

Figure B-2: Objective values of robust model at 10% variability as a function of Γ

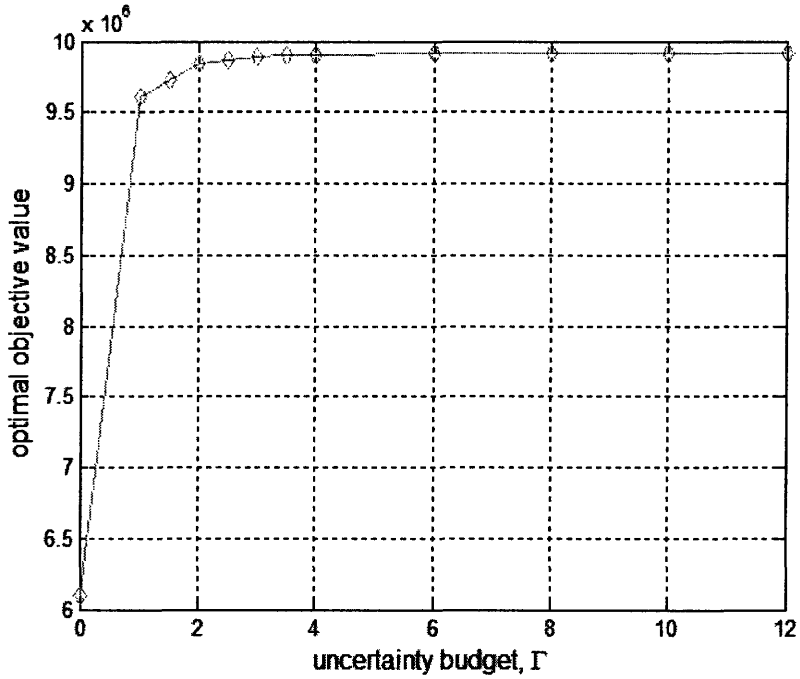

Figure B-3: Objective values of robust model at 20% variability as a function of Γ

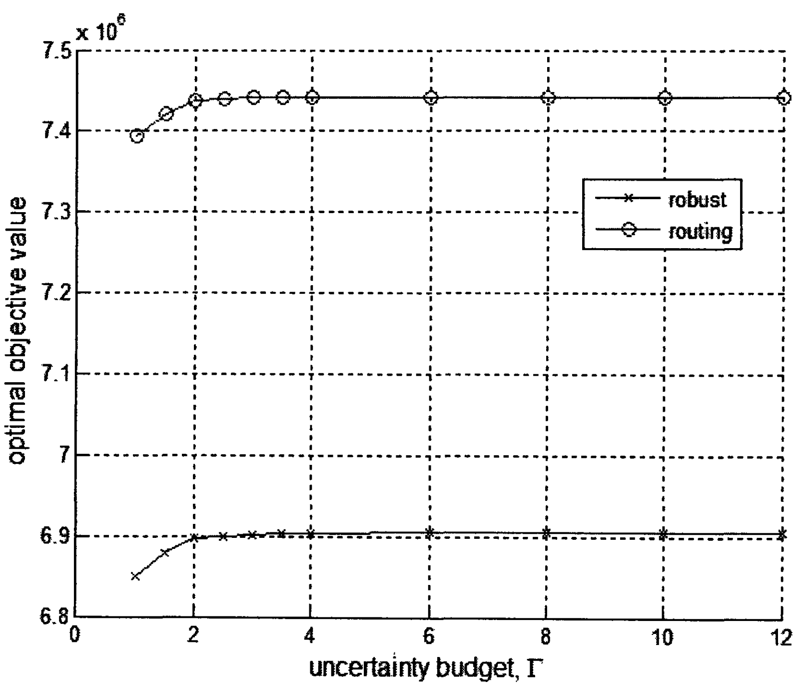

Figure B-4: Objective values of robust model at 5% variability as a function of Γ

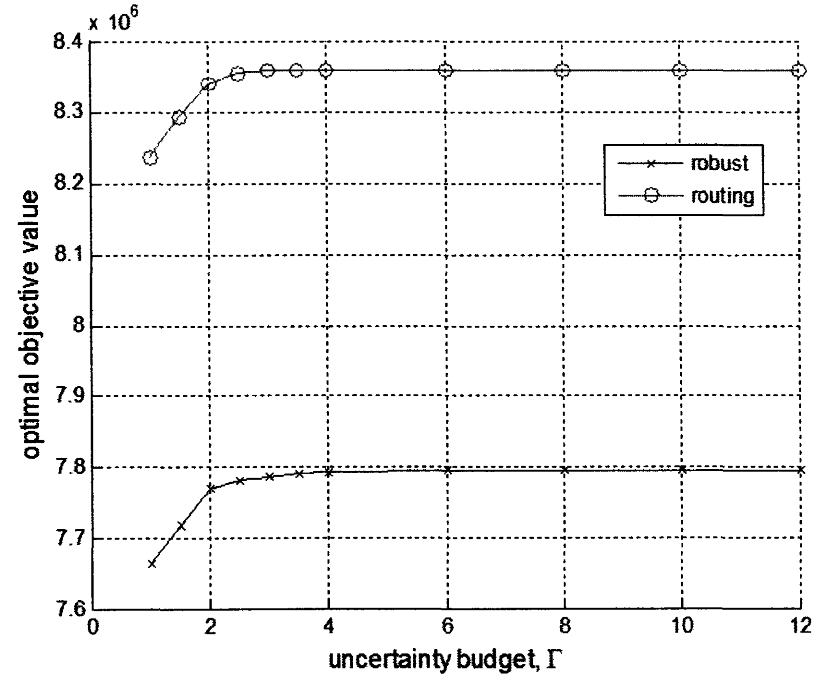

Figure B-5: Objective values of robust model at 10% variability as a function of Γ

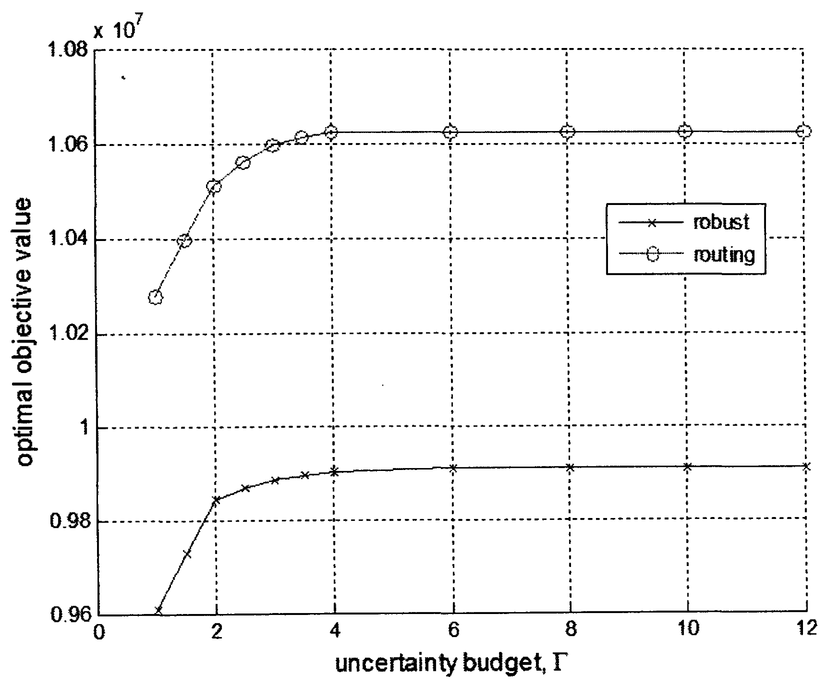

Figure B-6: Objective values of robust model at 20% variability as a function of Γ

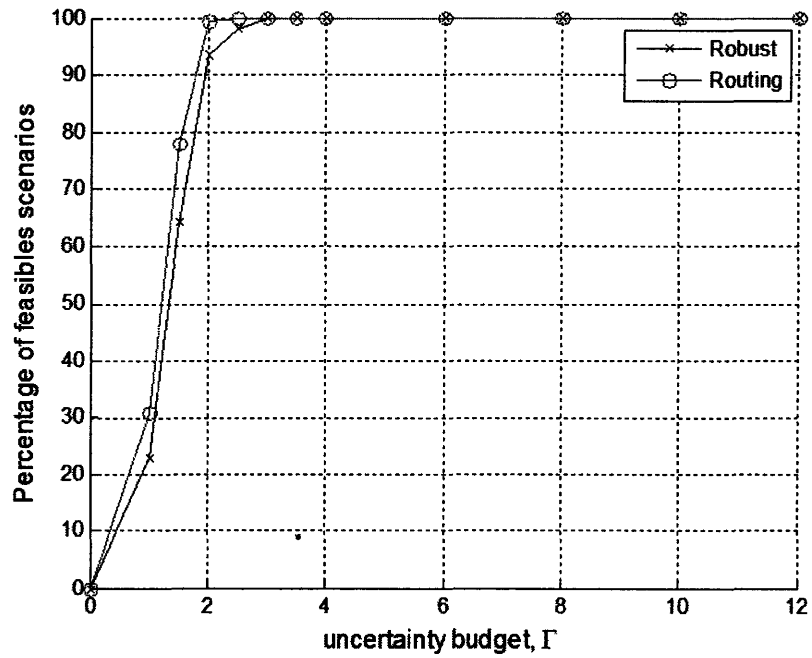

Figure B-7: Percentage of feasibile senarios at 5% variability as a function of Γ

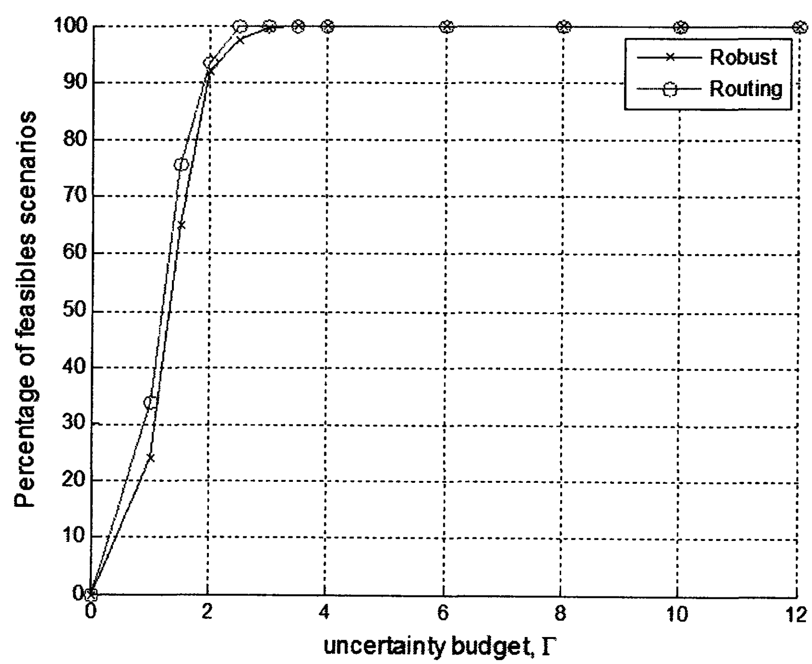

Figure B-8: Percentage of feasibile senarios at 10% variability as a function of Γ

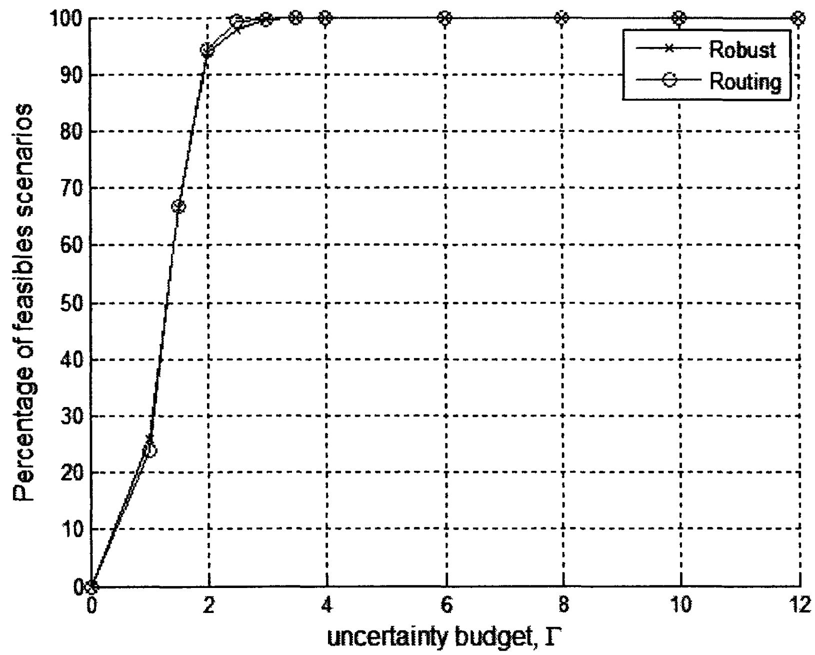

Figure B-9: Percentage of feasibile senarios at 20% variability as a function of Γ

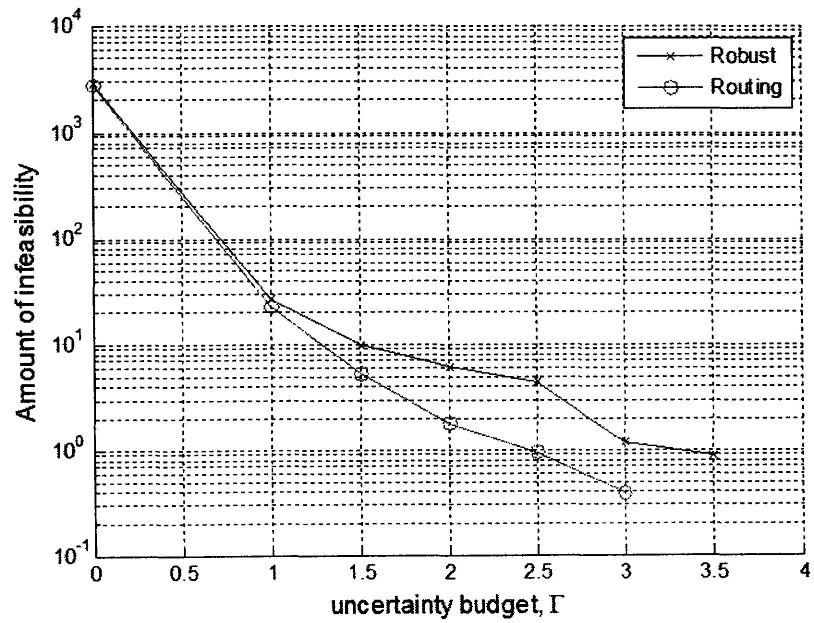

Figure B-10: Average infeasibility of senarios at 5% variability as a function of Γ

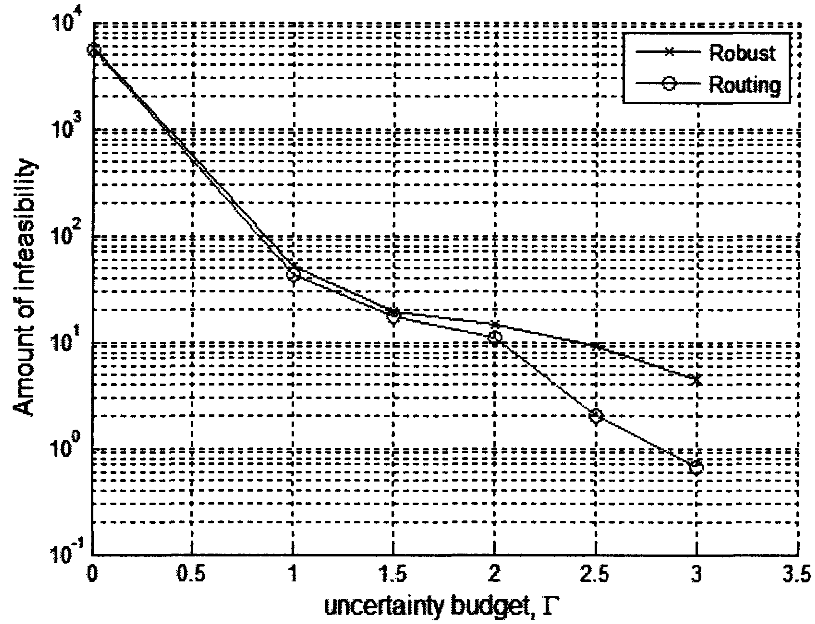

Figure B-11: Average infeasibility of senarios at 10% variability as a function of Γ

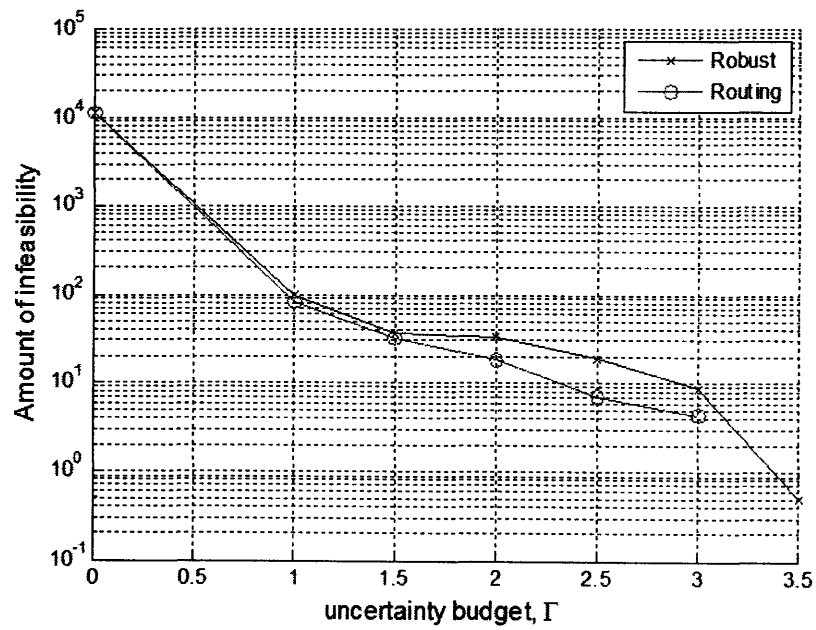

Figure B-12: Average infeasibility of senarios at 20% variability as a function of Γ

Bibliography

[1] Aharon Ben-Tal and Arkadi Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23:769–805, 1998.

[2] Aharon Ben-Tal and Arkadi Nemirovski. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25:1–13, 1999.

[3] Aharon Ben-Tal and Arkadi Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88:411–424, 2000.

[4] Aharon Ben-Tal and Arkadi Nemirovski. Robust optimization methodology and applications. *Mathematical Programming*, 92:453–480, 2002.

[5] Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations Reseach*, 52:35–53, 2003.

[6] Dimitris Bertsimas and Melvyn Sim. Robust discrete optimization and network flows. *Mathematical Programming*, 98:43–71, 2004.

[7] Daniel Bienstock and Nuri S. Ozbay. Computing robust basestock levels. *Discrete Optimization*, 5:389–414, 2008.

[8] I. Carrasco and Jorge R. Vera. Coorinación por descompositión de benders en cadena de abastecimiento. *Revista del Instituto Chileno de Investigacion Operativa(ICHIO)*, 2005.

[9] Miroslav Chlebik and Janka Chlebikova. Approximation hardness of the steiner tree problem on graphs. *Proceedings of 8th Scandinavian Workshop on Algorithm Theory (SWAT)*, pages 170–179, 2002.

[10] Rafael Epstein, Ramiro Morales, Jorge Seron, and Andrés Weintraub. Use of OR systems in the Chilean forest industries. *Interfaces*, 29:7–29, 1999.

[11] Joseph Geunes, Panos M. Pardalos, and Edwin H. Romeijn. *Supply Chain Management: Models, Applications, and Research Directions*, chapter 13. Springer, 2002.

[12] Laurent El Ghaoui and Hervé Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18:1035–1064, 1997.

[13] Laurent El Ghaoui and Hervé Lebret. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9:33–52, 1998.

[14] J. Karlsson, M. Rönnqvist, and J. Bergström. Short-term harvest planning including scheduling of harvest crews. *International Transactions in Operational Research*, 10:413–431, 2003.

[15] Andres Diaz Legues, Jacques A. Ferland, Celso C. Ribero, Jorge R. Vera, and Andrés Weintraub. A tabu search approach for solving a difficult forest harvesting machine location problem. *European Journal of Operational Research*, 179:788–805, 2007.

[16] Sergio Maturana, Fernando Ordóñez, Alfonso Perez, and Jorge R. Vera. Robust optimization: A case in forest operations planning. submitted to European Journal of Operational Research 2007.

[17] David Simchi-Levi, Xin Chen, and Julien Bramel. *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management.* Springer, 2005.

[18] Allen L. Soyster. Convex programming with set inclusive constraints and applications to inexact linear programming. *Operations Research*, 32:1154–1157, 1972.

[19] Jorge R. Vera, Andrés Weintraub, Monique Guignard, and Francisco Barahona. A lagrangian relaxation approach for a machinery location problem in forest harvesting. *Pesquisa Operacional*, 23:111–128, 2003.