

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600



Université d'Ottawa • University of Ottawa

**Analyses of Approximate Optimization
of Logistic Problems**

by

© Adel Bessadok

Masters of Science thesis
Systems Science programme
University of Ottawa

July 1996



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-21984-4

Abstract

The thesis focuses on methods of mathematical optimization for logistic planning with applications to the location of facilities and inventory of products. Algebraic models of location on networks are presented in increasing order of complexity and realism, emphasizing the presence of elemental models as building blocks of larger logistic systems. The thesis examines two small, well-known logistic problem formulations. For both applications, optimization programs are presented for which a classical solution approach is to find approximate optima because the representation encapsulates either combinatorial or functional complexities. Mathematical expressions of these approximations enable the calculation of their variations with the model parameters. Thus, constraint specifications of mathematical programs ease both the analysis of fast approximation and the analysis of variation of the optimal value as a function of each parameter of the model.

Starting with the uncapacitated facility location model, a simple heuristic for the location of facilities is compared with previous heuristics and exact algorithms and shown to yield an acceptable level of accuracy with respect to previous measures of quality.

In a study of capacity planning patterned after a traditional model of inventory control, a (Q,r) inventory system is analysed. The underlying mathematical model serves as a base for sensitivity analysis. As in the previous chapter, an approximation yields sufficient insight to predict the variation of optimal policies under varying conditions.

Acknowledgements

I gratefully acknowledge the financial support of the Mission universitaire et éducative de Tunisie for my studies at the University of Ottawa and the Operational Research Division (DCDS/DGOR) of the Department of National Defence which provided impetus for the accomplishment of this thesis.

I wish to thank Professors J. Couillard, F. Julien and J.-M. Thizy for allowing me to adapt some figures, tables and comments from their notes for the course in Operations Management at the Faculty of Administration of the University of Ottawa. Special thanks go to my supervisor, Professor J.-M. Thizy for his help, support and encouragement throughout my studies at the University of Ottawa. It was not practical to reiterate detailed credit for the sources making up the tight blend of my comparative analysis and accordingly, I waive all claims of copyright; readers interested in excerpting parts of the thesis should ask my supervisor for directions on citing original authorship properly.

To my parents.

This thesis was partially funded by NSERC grant #OGP 0042197 and National Defence Contract WO141-95-AA01.

List of Figures

Chapter 1	<u>Page</u>
Figure 1: Echelon Distribution Structure	14
Figure 2: Distribution system with additional echelons	18
Chapter 3	
Figure 1: Location of a central facility	43
Figure 2: Location of homogeneous facilities	44
Figure 3: Repetitive fixed cost function	45
Figure 4: A fixed order model for constant demand	49
Figure 5: Cost Components of the EOQ Model	50
Figure 6: Variation of Inventory in a (Q,r) System	54

List of Tables

	Page
 Chapter 1	
Table 1: Comparative Results for Some 33 x 33 Problems	23
Table 2: Comparative Results for Some 57 x 57 Problems	24
Table 3: Comparative Results for Some 100 x 100 Problems	25
 Chapter 2	
Table 1: Percentage values of the bound of Property 1	35
Table 2: Percentage values of the bound of Corollary 2 for standard data sets	36
Table 3: Percentage Values of the bound of Corollary 2 for complete Euclidian graphs	37
Table 4: Percentage values of the bound of Corollary 2 for trees with unit costs	37
Table 5: Percentage values of the bound of Corollary 2 for Euclidian trees	38
Table 6: Percentage values of the bound of Corollary 2 for sparse graphs with unit cost	38
Table 7: Percentage values of the bound of Corollary 2 for sparse Euclidian graphs	39
Table 8: Percentage values of the bound of Corollary 2 for random costs	40
 Chapter 3	
Table 1: Variation of Inventory Cost with Order Quantity	52
Table 2: Random sample of shortages	55
Table 3: Tally of order sizes	55
Table 4: Cumulative frequency of stock availability	56
Table 5: Expected shortage	57

Table of Contents

Introduction 1

Chapter 1

Mathematical programming formulation of multi-stage multi-commodity distribution systems .. 6

- 1. Mathematical formulations of distribution systems 7
 - 1.1 The Uncapacitated Facility Location Problem 8
 - 1.2 The Capacitated Facility Location Problem 11
 - 1.3 Multi-echelon distribution systems 13
 - 1.4 Flexible formulation and solution of distribution models 17
- 2. Solution methods for the Uncapacitated Facility Location Problem 20
 - 2.1 Heuristic algorithms 21
 - 2.2 Performance of heuristics versus exact algorithms 22
 - 2.3 Efficient optimal algorithms 24
 - 2.4 The Challenge of Large Uncapacitated Facility Location Problems 26

Chapter 2

Analysis of a simple heuristic for the uncapacitated facility location problem 29

- 1. The greedy heuristic for the uncapacitated facility location problem 30
- 2. A new bound for the greedy location heuristic 32
- 3. A disaggregation of the uncapacitated facility location problem 33
- 4. Computational results 35

Chapter 3

Aggregate and average network capacity management 41

1. **Homogeneous capacity planning 41**

2. **Repetitive fixed charge problem: batch manufacturing 47**

 2.1 **Review of Inventory Control 47**

 2.2 **A Fixed Order Quantity Model for Deterministic Demand 48**

3 **Distribution network with stochastic Demand 53**

 3.1 **Fixed Order Quantity Models for Stochastic Demand 54**

 3.1.1 **Service level based on frequency 56**

 3.1.2 **Service level based on expected unsatisfied demand 57**

 3.2 **Sensitivity Analysis of Continuous Inventory Control with
 Service Level Constraint 59**

 3.2.1 **Mathematical programming formulation 60**

 3.2.2 **Sensitivity analysis 63**

Conclusion 72

Bibliography 74

Introduction

Proper location and management of facilities by service and manufacturing firms is a crucial part of logistic planning, as it affects their responsiveness, flexibility, and eventually their long-term profitability. At the operational level, widespread decreases in corporate inventories have highlighted the importance of judicious locations, not only as a cost-cutting measure, but as a vital link in the design of synchronized distribution. For the supply of products, the problem is compounded by the rising cost of dedicated equipment of the handling facilities.

With this particular focus, the thesis will examine various methods of mathematical optimization for logistic planning with highlight on the location of facilities and capacity planning. The vast majority of traditional models considers the distribution network and its facilities as given, and determine their required capacities. This approach subsists at a time when both internal distribution services and logistic contractors seek to consolidate and re-size their facilities. Thus, to capture current requirements, the design of distribution systems must combine distribution (transportation) and capacity management, devoting more attention to detailed operations such as inventory control. Each discipline commands a vast literature with frequent opportunities for overlaps. Yet, few models achieve a balanced integration [Albright, 1989; Cohen, Kleindorfer and Lee, 1989; Graves, 1988; Lee and Moinzadeh, 1989; Jack, 1992; Pyke, 1990].

Practitioners' approach to selecting sites and managing facilities is often reduced to a simple scenario analysis with the assumption that fixed costs, equipment and manpower capacities are predetermined by the operations of other equipment. For long-term planning [Yip, 1995], the shortcoming of the approach is clear. Another common practice is to adapt time-honoured models such as the transportation problem or the economic order quantity to the situation at hand - an approach prone to obvious weaknesses - or to customize them at much cost and delay. Finally, a classical suggestion is to design an overall logistic system that will coordinate maintenance, inventory and distribution activities. This approach generally leads to large formulations manned by specialists, which absorb and provide great amounts of data, but often fail to grasp some important idiosyncrasies of the given

situation. In contrast, the thesis emphasizes small, well-known formulations as building blocks of larger logistic systems.

A practical motivation for this thesis was the design of logistic networks as used for Canadian national defence, which are typically large and complex [Nemhauser, 1994; Wymore, 1993]. In response, the Directorate of Logistics Analysis of the Department of National Defence (DND) has developed various modelling systems, for instance the simulation of inventory management which allows a multi-echeloned, multi-indentured inventory problem to be investigated dynamically [Martel, 1991]. However, field managers find the models increasingly difficult to use under the fast change of distribution modes caused by evolving technology and budgetary pressures. Following a national trend toward direct sourcing, smaller inventories are kept at fewer locations, but suppliers are allowed to intervene more flexibly. The adequacy of the location of supplies is no longer shielded by costly inventory or maintenance equipment, and previous decision software using central, multi-stage replenishment proves often inadequate. In this context, the design of decision systems for level of repair analysis was dominated by themes of location, capacity planning and inventory control which are analysed in this thesis.

The analytical formulation of such models gives an opportunity to review their role in exact and approximate optimization. Interestingly, the problems presented in this thesis are not only related by their practical importance, but by theoretical links as well. Optimization programs are presented for which a classical solution approach is to find approximate optima because the representation encapsulates either combinatorial or functional complexities.

In a first chapter that introduces locational analysis in the distribution network, various approximations of a simple NP-hard model are reviewed. In a second chapter, various measures of quality of a simple approximation are compared in view of real-world data sets.

In the last chapter, a study of capacity planning borrows from the logistic models of deterministic and stochastic inventory control; traditional approximate expressions are replaced by a mathematical programming formulation, which enables the calculation of its variation with the model parameters.

Mathematical models of logistic optimization

Mathematical models answer two types of requirements which often partition them into two (overlapping) classes:

- represent a process (or at least the decision-maker's view of the process) and thus refine and support a mental model;
- document the conditions in which the process takes place, usually with relevant data, and thus offer the base of a module of a large decision support system.

Algebraic optimization models can also fulfil a dual role:

- numerical specification: they specify the conditions that a numerical optimization must satisfy; in this sense, they are translated into data structures as input of an external solution algorithm.
- functional representation: by displaying relationships between the entities of the model, they document it and facilitate its reuse. Many optimization models are NP-hard or intractable [Garey, 1979] and therefore the numerical specification may not offer operational means of determining the value of the optimum. Yet, these specifications can be invaluable as they lend foundations for larger models.

Indeed, the optimal value of the two models presented in the thesis cannot be specified directly in algebraic, closed form. However, their constraint specifications ease both the analysis of fast approximation and the analysis of variation of the optimal value as a function of each parameter of the model.

Location models for distribution systems

Many rough-cut distribution planning systems (among which level-of-repair analysis motivating this thesis) take an aggregative approach to network modelling. The hypothesis is that each centre has identical characteristics such as repair time, level of inventory, speed of repair. In fact, these levels vary and should be taken into account, especially in view of two recent trends:

- the increased responsibility of the decentralized management emphasizes the autonomy and uniqueness of each centre, and favours local initiative. Reducing each centre to a common denominator forecloses its ability to display natural advantages;
- in a budget curtailing situation, many institutions such as the Armed Forces reassess the number of facilities devoted to the support of their operations. In particular, they must select which centre will be closed. Therefore, they need to differentiate between the varying efficiency or natural advantage of each site. Accordingly, accepted methods of location analysis enable decision-makers to select or reject sites for operations support. Along the distribution network, quantitative models described in the first chapter of the thesis determine the sites and pattern of commodity flows selected.

In this vein, Chapter 1 presents algebraic models of location, in increasing order of complexity and realism, emphasizing the presence of an elementary block known as the (un)capacitated facility location model. A brief review of classical solution methods addresses the following questions:

- when can the problem be solved optimally?
- how well do simple approximate methods perform?
- for which data configuration is either approach best suited?
- what are the costs and benefits of special-purpose algorithms?

In particular, a simple heuristic is compared with previous heuristics and exact algorithms. A new measure of accuracy is reviewed in the next chapter and compared with previous measures of quality.

Aggregate and average network capacity management and inventory models

Stochastic programming methods have been applied to the design of logistic networks under fluctuating demand. Within the focus of the thesis, i.e., assessing the potential contribution of well-

known models, standard techniques used for inventory control for stochastic demand are shown to have a double relevance to logistic management:

- offer a general approach to repetitive fixed charge problems as occur in capacity planning,
- present immediate interest to logistic design which must integrate issues of feasibility and cost raised by recurring operations.

Inventory management, often viewed as an unrewarding routine, determines critically the efficiency of logistic operations. Yet, in practice, warehouse managers use rules-of-thumb, experience, to reorder parts so as to maintain a steady inventory level at their facility. However, the level of inventory itself is rarely debated, and inventory replenishment practices tend to inflate their size of inventory, since managers attempt to avoid stockout for which they may be directly penalized. Hence, inventory can reach unrealistic proportions. This has been particularly acute for National Defence systems, leading to public scrutiny of inventory levels [Pugliese, 1996].

This last chapter adapts a traditional (Q,r) inventory control system to the need for scenario analysis. The underlying mathematical model serves as a base for sensitivity analysis. As in the previous chapter, an exact measure of variation is approximated, yielding sufficient insight to predict the variation of optimal policies under varying conditions.

Chapter 1

Mathematical programming formulation of multi-stage multi-commodity distribution systems

The design of distribution systems appears to be a well defined endeavour with a clear goal (the efficient supply of some commodities), a common set of solution techniques, and a large number of applications of theoretical and practical interest. This apparent unity can conceal a wide variety of situations, from the design of the physical distribution of products, to the location of emergency services. In their most general definition, the goods can be some units of (dis)utility, or percentages of logical requirements, and their distribution accordingly be highly ethereal.

A vast number of such logistic problems revolve around physical movement hinging on the choice of sites to locate some facilities [Arcelus, 1989]. They occur frequently in distribution or planning systems such as:

- warehouse and plant location [Simkin, 1989],
- air, railroad and mass transportation,
- design of telephone [Kochman and McCallum, 1981; Mirzaian, 1985], television and computer networks [Boffey, 1989], siting of broadcasting stations [Drezner, 1988],
- layout of electric power systems,
- construction of oil and gas pipelines,
- planning of health centres, public services, police and fire stations,
- location of lock-boxes, collection centres.

Such situations also arise in related managerial problems: production planning [Krarup and Bilde, 1977; Pochet and Wolsey, 1988], machine replacement [Yano, 1985 (U. Mich)], vehicle routing [Magnanti, 1981], homogeneous clustering [Mulvey, 1982], capital budgeting, aggregation of

database information, optimal constitution of portfolios, or in seemingly unrelated questions such as the clinical detection of glaucoma [Kolesar, 1980].

It is not surprising that modelling distribution systems can have quite an expansible scope. This thesis assesses another source of heterogeneity: the variety of formulations and optimization methods for solving distribution problems. For example, consider the location of regional warehouses to serve a given set of customers, where each warehouse has a maximum capacity and all the distribution costs are predetermined. Many companies require each of their customers to order from one regional warehouse only, for contractual or organizational simplicity. The opposite choice, to let each customer be supplied from several warehouses, may well generate enough savings to be attractive to the manager. But such a seemingly innocuous modification of the problem corresponds to markedly different formulations and algorithmic treatments.

1. Mathematical formulations of distribution systems

The characteristics of locational problems can be extremely diverse: the distribution may involve several products or several stages. The economic criterion may vary from the minimization of the distribution cost, to the minimization of the penalty incurred by the least favoured area serviced, the aggregation of each user's satisfaction [Beckman, 1956] or dissatisfaction [Drezner, 1989]. Concomitant or conflicting criteria may be at stake [Bhaskaran, 1990; Current, 1990; Aly, 1990]. The demand for services may be stochastic (location of fire stations) [Mirchandani, 1983; Louveaux, 1993] or deterministic, static or dynamic: in particular, the choice of a location for a plant must be made, bearing in mind possibilities of expansion. A survey of dynamic location models has been published in [Erlenkotter, 1975; Megiddo, 1986]. For some of these models, the optimizing goal itself becomes elusive. Besides the profuse bibliography of [Lea, 1973], some surveys particularly relevant to the readers are listed below in chronological order: [Balinski, 1969; ReVelle, 1970; Eilon, 1971; Hansen, 1972; Elshafei, 1974; Francis and Goldstein, 1974; Francis et White, 1974; Kaufman, 1975; Geoffrion, 1975; Guignard, 1977; Salkin, 1975; Jacobsen, 1976; Cornuejols, 1977; Krarup, 1977, 1979; Erlenkotter, 1975; Wong, 1978; Handler, 1979; Van Roy, 1980; Aikens, 1982, 1985; Francis

1983; Love 1988; Brandeau, 1989; Rushton, 1989; Schilling, 1993; Ghosh, 1993]. Among books on location, some are dedicated to special topics, e.g. discrete location problems, [Francis, 1974 (P.H.), 1984; Mirchandani, 1990; Daskin, 1995]. The host of methods proposed rarely addresses the problem facing the decision-maker. Fortunately (owing to its physical nature), the problem at hand can often be decomposed or simplified into common models.

Following are various distribution models, presented in order of complexity of formulation, the latter superposing the structures of several simpler problems. Note that, even if some problems are harder to formulate and may demand a larger, more sophisticated solution algorithm, they may not be inherently harder to solve than some problems which can be formulated very concisely: for example, the large number of constraints may well reduce the number of feasible solutions of the problem and therefore its difficulty.

At the base, the uncapacitated facility location problem (UFLP) and the capacitated facility location problem (CFLP) constitute the simplest building blocks of the varied models proposed. They are discrete, static, deterministic, one-product, one-stage models to minimize the distributor's costs.

1.1 The Uncapacitated Facility Location Problem

"The simple plant location problem is one of the simplest mixed integer problems which exhibits all the typical combinatorial difficulties of mixed (0,1) programming and at the same time has a structure that invites the application of various specialized techniques [Guignard, 1977]."

The simple plant location problem consists of selecting facilities to be open among a finite set J of potential sites in order to minimize the cost of serving a given set of customers. Since no capacity constraints on the sites are considered, this problem is also referred to as an uncapacitated facility location problem. A cost c_{ij} is incurred to satisfy the totality of the demand of Customer i from Location j if a facility is kept open there, which can be maintained at this site at a cost f_j .

The mathematical formulation of UFLP is then:

$$z = \min \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} f_j y_j, \quad (1)$$

$$\sum_{j \in J} x_{ij} = 1 \quad \text{all } i \in I, \quad (2)$$

$$0 \leq x_{ij} \leq y_j \quad \text{all } i \in I, j \in J, \quad (3)$$

$$y_j = 0 \text{ or } 1 \quad \text{all } j \in J. \quad (4)$$

where the variables x_{ij} indicate the portion of the demand of Customer i satisfied by Location j and y_j is 1 or 0 according as Site j is selected or not.

UFLP occupies a central position among popular variants for which solution methods overlap each other to a considerable extent; their similarities are emphasized in a unified introduction.

Probably the most common additional requirement is to select exactly p facilities [Cornuéjols, 1978], which means the addition of the following constraint:

$$\sum_{j \in J} y_j = p \quad (5)$$

The resulting problem is called the p -facility location problem [Cornuéjols, 1978]. In the thesis, this term will be typically subsumed by the better known p -median problem [Beasley, 1984; Captivo, 1991; Labbe, 1992], in which the costs are Euclidian distances and there is no fixed charge f_j . A common subcase of (5) arises when, for $p-1$ additional facilities indexed by $j = 0, -1, -2, \dots, 2-p$, the fixed charges $f_j = 0$ and the costs c_{ij} are very large, which practically enforces only an upper limit on the number of other facilities to be chosen [Cornuéjols, 1977]:

$$\sum_{j \in J} y_j \leq p \quad (6)$$

The addition of the requirement (6) generalizes the UFLP, because it can be waived by setting the upper limit p equal to the total number of potential sites. We will refer to p -median problem in a generic fashion, whether equality is enforced in (6) or not. Chapter 2 of the thesis focuses on UFLP; the other models presented in the following sections are not introduced in view of direct implementation, but to assess the potential of UFLP in logistic modelling. Optimal algorithms to solve the UFLP or p -median problems have followed three basic strands during two decades of successive improvements.

- Early algorithms were designed as direct applications of general optimization techniques such as branch-and-bound [Efroymsen, 1966], Benders' or Dantzig-Wolfe decomposition [Balinski, 1963, Swain, 1974], group-theoretic approach [Garfinkel, 1974; Dearing, 1992], that can be improved by heuristic choices in the enumeration [Spielberg, 1969; Khumawala, 1972, 1973, 1974; Jarvinen, 1972; Hansen, 1972; El-Shaieb, 1973].
- The specific structure of the UFLP and its relatives favours the use of streamlined algorithms using special-purpose data structures: [Marsten, 1972] for the p -median problem and [Schrage, 1975] for the UFLP.
- The dual formulations of the problem introduced by [Diehr, 1972 and Geoffrion, 1974] were implemented by fast algorithms [Gao, 1992], such as those based on Lagrangian relaxation [Galvao, 1993; Mulvey, 1979], subgradient optimization [Neebe, 1975; Narula, 1977; Christofides, 1982; Cornuejols, 1977], non-linear optimization [Conn, 1990] or on local ascent search [Bilde, 1977; Galvao, 1980] for the p -median problem.

The best known representation of this family is the dual adjustment procedure [Erlenkotter, 1978], described by Erlenkotter's own characterization of his algorithm:

... this solution approach may be close to the ultimate in efficiency for the problems examined.

1.2 The Capacitated Facility Location Problem

In distribution systems, each potential distribution centre may have a capacity s_j which can be incorporated into the previous model by introducing Constraints (9), thus forming the Capacitated Facility Location Problem (CFLP):

$$z = \min \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} f_j y_j, \quad (7)$$

$$\sum_{j \in J} x_{ij} = d_i \quad \text{all } i \in I, \quad (8)$$

$$\sum_{i \in I} x_{ij} \leq s_j y_j \quad \text{all } j \in J, \quad (9)$$

$$0 \leq x_{ij} \quad \text{all } i \in I, j \in J, \quad (10)$$

$$y_j = 0 \text{ or } 1 \quad \text{all } j \in J. \quad (11)$$

where, in addition to preceding definitions given for UFLP:

- c_{ij} is the unit cost of satisfying the demand of Customer i from Location j ,
- d_i is the demand of Customer i ,
- x_{ij} is the amount of demand of Customer i satisfied by Facility j .

The solution methods parallel those of the UFLP: enumeration [Davis, 1969, Ellwein, 1971], directed by the greedy [Salkin, 1975] or other heuristics [Spielberg, 1970; Khumawala, 1974, 1977; Sridharan, 1993], Lagrangean relaxation [Geoffrion, 1978; Beasley, 1988] or dual adjustment [Guignard, 1979]. The fastest algorithm to solve the capacitated plant location problem [Van Roy, 1980, 1986] is based on the following remark [Van Roy, 1980]:

"..., experience in practice [Van Roy et al., 1981] has indicated that capacity constraints for new facilities tend to be weak since capacity is yet to be established. Existing facilities, though critically capacitated, do not create duality gaps since the corresponding y_j 's are fixed to one. As such, relaxing capacity constraints is a natural approach."

Instead of enforcing the capacity constraints (9), the method of Lagrangean relaxation uses a vector of multipliers w to price the amount by which the constraints are satisfied or violated. The objective function becomes:

$$z(w) = \min \sum_{i \in I} \sum_{j \in J} (c_{ij} + w_j) x_{ij} + \sum_{j \in J} (f_j - w_j s_j) y_j \quad (9.w)$$

Different sets of prices w are tried systematically; the Lagrangian dual is then defined as:

$$z_L = \max_{w \geq 0} z(w).$$

In general, $z > z_L$ and z_L itself is approximated by solving a series of problems with a modified objective. The procedure hinges therefore on the existence of a fast algorithm to solve each such subproblem, a UFLP, for which the variable distribution costs may no longer satisfy the original properties of the physical system (e.g. positive fixed costs or symmetric variable costs that may reflect Euclidian distances or satisfy the triangular inequality), thereby disqualifying many special-purpose algorithms which would solve the original uncapacitated version of the problem, in particular most algorithms for the p -median problem. A different algorithmic approach is based on the polyhedral characterization of the polytope of feasible solutions [Aardal, 1996], a method pioneered by the UFLP, as reviewed in Section 2.3 .

The next section present a more realistic, yet complex distribution model, using CFLP and therefore UFLP as building blocks.

1.3 Multi-echelon distribution systems

In many problems, several commodities are in demand [Crainic, 1989; Markis, 1989]. In a simple, but well-known multi-stage multi-commodity distribution system [Geoffrion et al., 1974], several commodities can be processed by several plants within predetermined production capacities. Each commodity may be required at several customer zones, to be provided in deliveries shipped via regional distribution centres (abbreviated DC). Each customer zone can receive shipments from at most one DC. The allowable total annual throughput of each DC may be constrained by lower and upper bounds. The DC costs are decomposed as variable linear costs and fixed costs incurred upon making use of the DC's facility. Transportation costs are assumed to be simply linear.

To meet the given demands at minimum total distribution cost, one must determine which DC sites should be used, the sizes of the DC, which customer zones each DC should serve and what transportation flows should be chosen for all commodities, as shown in Figure 1. The mathematical formulation of the problem uses the following notation.

$i \in I$	index and set of commodities,
$j \in J$	index and set of plants,
$k \in K$	index and set of possible distribution centre (DC) sites,
$h \in H$	index and set of customer demand zones,
s_j	available supply for Commodity i at Plant j ,
d_{ih}	demand for Commodity i in customer Zone h ,
$\underline{V}_k, \bar{V}_k$	lower and upper bounds on the total annual throughput for a DC at Site k ,
f_k	fixed cost of operating a DC at Site k ,
v_k	variable unit cost of flow through a DC at Site k ,
c_{ijkh}	average unit cost of processing Commodity i from Plant j through DC k to customer Zone h ,
x_{ijkh}	a variable representing the flow of Commodity i shipped from Plant j through DC k to customer Zone h ,

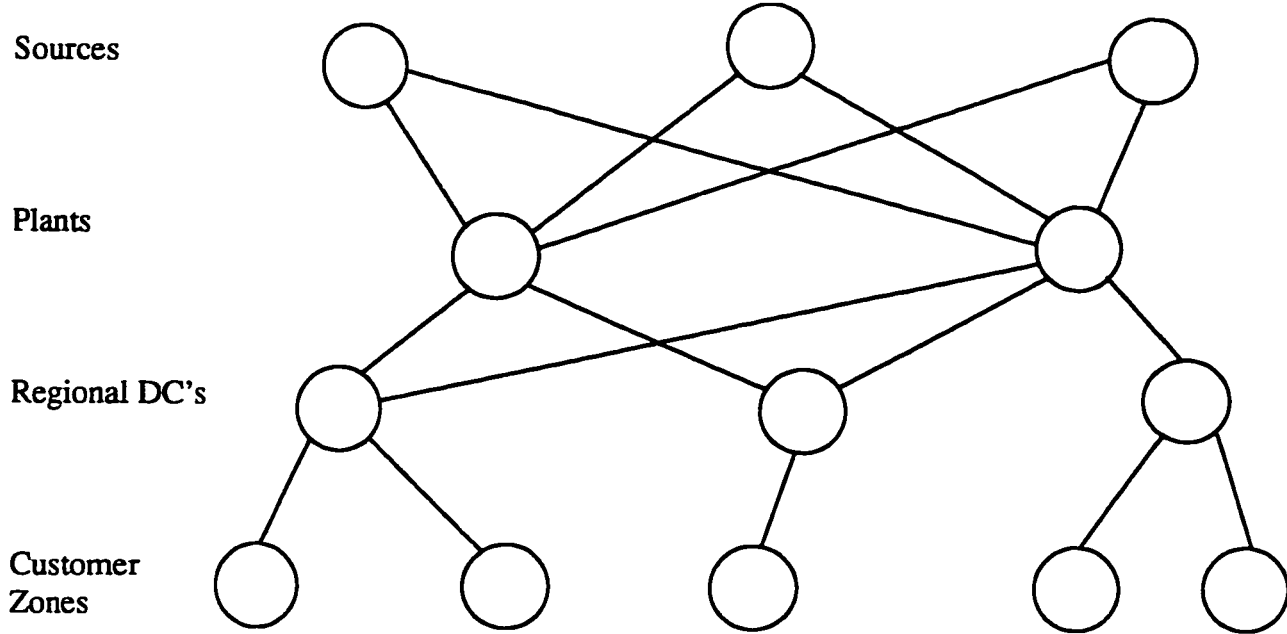


Figure 1: Echelon Distribution Structure

y_{kh} a variable that is 1 if DC k serves customer Zone h and 0 otherwise,
 z_k a variable that is 1 if a DC is located at Site k and 0 otherwise.

The problem can be formulated as the following program.

$$\text{Minimize } \sum_{ijkh} c_{ijkh} x_{ijkh} + \sum_k [f_k z_k + v_k \sum_{ih} d_{ih} y_{kh}] \quad (12)$$

$$\text{subject to } \sum_{kh} x_{ijkh} \leq s_{ij} \quad \text{all } i \in I, j \in J \quad (13)$$

$$\sum_j x_{ijkh} = d_{ih} y_{kh} \quad \text{all } i \in I, k \in K, h \in H \quad (14)$$

$$\sum_k y_{kh} = 1 \quad \text{all } h \in H \quad (15)$$

$$\underline{V}_k z_k \leq \sum_{ih} d_{ih} y_{kh} \leq \bar{V}_k z_k \quad \text{all } k \in K \quad (16)$$

$$x_{ijkh} \geq 0 \quad \text{all } i \in I, j \in J, k \in K, h \in H \quad (17)$$

$$y_{kh}, z_k = 0, 1 \quad \text{all } h \in H, k \in K \quad (18)$$

In this model:

- Constraints (13) limit the total shipment of Commodity i out of Plant j to the supply of this plant,
- Constraints (14) require that demand be met when $y_{kh} = 1$ and that x_{ijkh} be 0 for all $i \in I$ and $j \in J$ when $y_{kh} = 0$,
- Constraints (15) ensure that each customer zone be served by exactly one DC,
- Constraints (16) keep the total annual throughput of DC k between \underline{V}_k and \bar{V}_k , which corresponds to enforcing $z_k = 1$, and prevent any shipment through DC k when $z_k = 0$.

The double stage structure of the model is easily discernible in the mathematical formulation. Denote $|I|$ the cardinality of Set I . For the first stage (the plant-DC distribution), the subproblem (12)-(14) is reminiscent of a capacitated plant location problem, itself decomposable into $|I|$ capacitated plant location problems P_i , one for each of the $|I|$ commodities. To illustrate the analogy better, the index i of each problem P_i will be dropped: the customer index is j and each facility q of P_i corresponds to a distribution channel formerly indexed by DC k and Customer h of the original formulation (Q is therefore a subset of $K \times H$).

Problem (12)-(14) can then be reformulated as :

$$\min \sum_{j \in J} \sum_{q \in Q} c_{qj} x_{qj} + \sum_{q \in Q} f_q y_q \quad (19)$$

$$\sum_{q \in Q} x_{qj} \leq s_j \quad \text{all } j \in J \quad (20)$$

$$\sum_{j \in J} x_{qj} = d_q y_q \quad \text{all } q \in Q \quad (21)$$

$$x_{qj} \geq 0 \quad \text{all } j \in J, q \in Q$$

$$y_q = 0 \text{ or } 1 \quad \text{all } q \in Q$$

Some simplification was obtained by setting $f_q = f_k + v_k d_q$ where $d_q = d_h$ for every Commodity i , DC k , Market h and $q = (i, h) \in Q$. The second stage of the distribution, the DC-customer distribution, defines a subproblem (12), (15)-(16), also decomposable into $|I|$ capacitated location problems (with lower bounds on the volume of demand processed as well as capacities) in which the facility and customer indices are k and h respectively (with corresponding index sets J and H).

$$\min \sum_{k \in K} f_k z_k + \sum_{k \in K} \sum_{h \in H} g_{kh} y_{kh} \quad (22)$$

$$\sum_{k \in K} y_{kh} = 1 \quad \text{all } h \in H \quad (23)$$

$$\underline{V}_k z_k \leq \sum_{h \in H} d_h y_{kh} \leq \bar{V}_k z_k \quad \text{all } k \in K \quad (24)$$

$$y_{kh}, z_k = 0 \text{ or } 1 \quad \text{all } k \in K, h \in H$$

where $g_{kh} = v_k d_h$ for all $k \in K$ and $h \in H$.

1.4 Flexible formulation and solution of distribution models

Many real distribution environments adopt a distribution mode known as single-sourcing, where each customer must be served from only one facility, as in the preceding model, in order to promote a marketing policy, a simplified order processing system, local competitive or regulatory conditions. On the other hand, other distribution plans allow each customer to be supplied by several facilities, a configuration often called multi-sourcing. This alternative is often an inexpensive way to manage a distribution network, because of variations in the way facilities process the products and the resulting difference of production costs, capacity constraints, or other distribution criteria. To obtain a corresponding formulation, Constraints (18) can be relaxed to:

$$0 \leq y_{kh}, z_k \leq 1 \quad \text{all } j \in J, h \in H \quad (25)$$

A given variable y_{kh} represents actually the satisfaction from one facility of a given share of a customer's demand for all products, and a variable z_k enables a scaling of Facility k . The binary restriction (18) on the variables forced these factors to be 0 or 1, i.e., provisioning the customer from a single source or opening a fully-sized plant. Even under the relaxed constraints (19), almost all the variables y_{kh} are 0 or 1 because from each facility, the demand of at most one customer may be divided, this case occurring only if the facility's capacity becomes restricted. Hence, only a few variables y_{kh} may need to be forced to 0 or 1 and solution times remain reasonable. On the other hand, the variables z_k , modelling fixed charges for opening facilities, must usually be forced to assume values of 0 or 1. To achieve a single/multiple source combination, a mix of continuous and 0-1 variable requirements may obviously present a very flexible structure.

In the preceding formulation, the echelon structure is simple, and so is the treatment of the multiple products. On the other hand, if direct shipments from plants to Customers are allowed, each plant must be twinned with a dummy warehouse.

A simple extension of the two-echelon formulation can of course yield a three-echelon system. To give more flexibility to the regional distribution, an echelon of field DC's may be added as in Figure

2. In fact, many real world distribution systems involve more echelons, e.g., two levels of warehousing (e.g. national and regional); inbound shipments from vendors may also be added above the plant level. Integer variables can then be introduced either at the level of facilities (the plant echelon) to select their optimal location or at an intermediate level (a warehouse echelon) for the selection of warehouse facilities. Moreover, the model of [Geoffrion et al., 1974] can be remedied to allow alternative product to plant allocations and regional distribution centres. One method consists of including an upper echelon for product source and allowing some zero-one source-plant allocation variables. Another method is to create fictitious (product-plant) facilities and to open or close (product-plant) facilities. Both solutions correspond to a genuine two-stage distribution system.

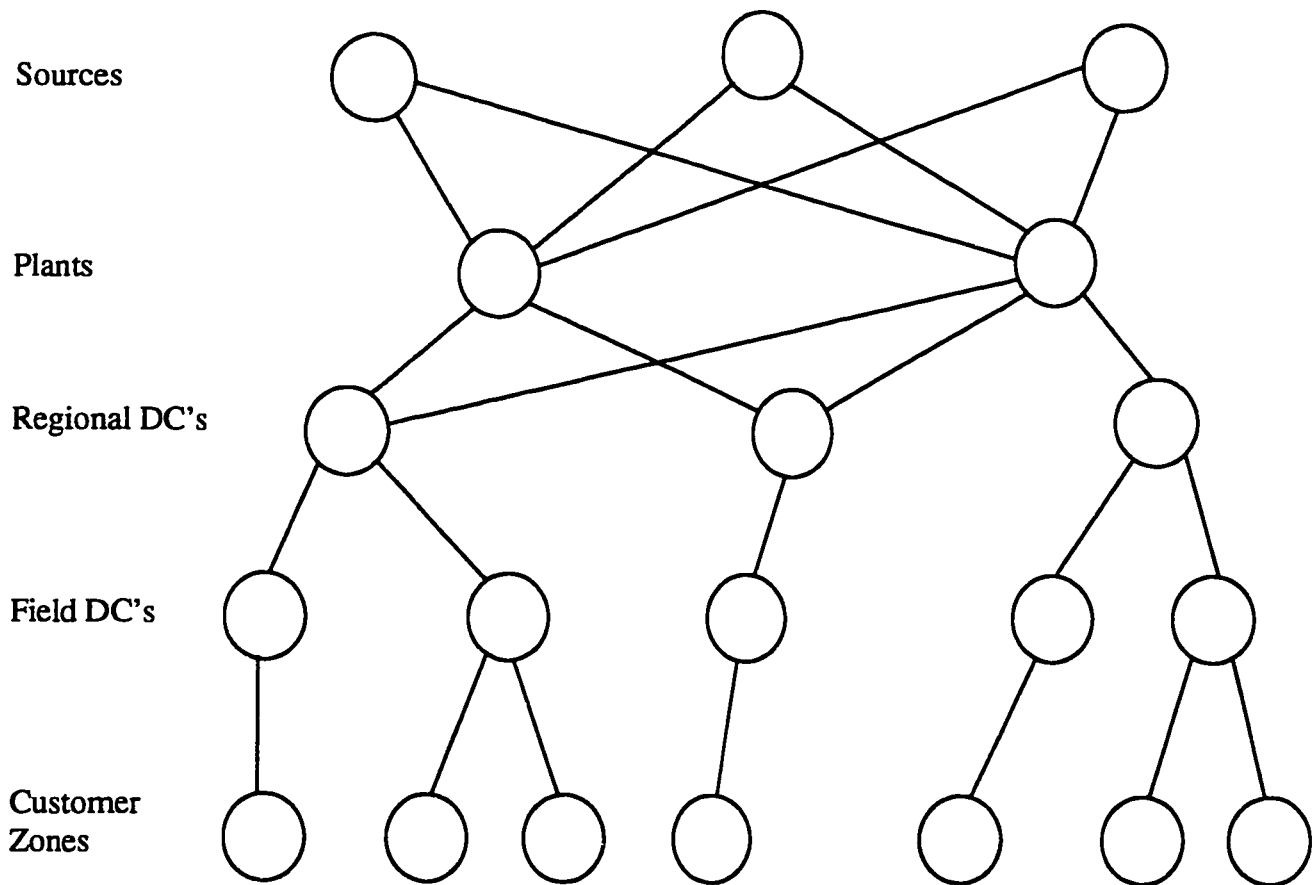


Figure 2: Distribution system with additional echelons

From a mathematical perspective, all binary requirements on variables are of the same nature. Yet, the practical difficulty of solving the model will hinge crucially on the number of plants, i.e. the number of variables z_k , whereas one can expect the computational burden to be proportional to the number of markets, rather than to the number of variables y_{kt} . This difference illustrates again the central question of adapting the model and its solution to the distribution system at hand: whereas, for a corporate model, the number of plants is not very large and the increase of problem size caused by the inclusion of plant location variables is moderate, this is certainly not the case of an independent retail distribution system (for example automotive spare parts). The choice of formulation raises by itself issues of system design. A succinct outline of a proper selection procedure of a given distribution system is listed below [Schutt, 1979]:

- a) choose the type of model (inventory problems may completely overshadow the distribution system);
- b) assess the relative sizes of the potential sets of products, plants, warehouses, customers;
- c) study the feasibility of regional decentralization (either total; or with a national master problem);
- d) determine whether an echelon structure (plant, DC, customer,...) is appropriate or too restrictive with respect to allowances for plant and warehouse transshipments or direct deliveries from the factory;
- e) decide where the orders can be split between several origins and the a-priori economic importance of such a distribution policy;
- f) investigate the potential bottlenecks (tight capacities), bearing in mind that Lagrangian techniques are less useful where the relevant constraints are binding;

- g) state the real objective of the design: importance of the near optimal solutions and of sensitivity analysis.

Finally, as in all modelling activities, a first task is to question the relevance and accuracy of the data.

The formulations resulting from such flexible models can also be solved efficiently by decomposition methods. The simplest decomposition distinguishes location (zero-one) from allocation (continuous) variables. A multi-stage decomposition algorithm paralleling the possible multi-stage structure offers the intuitive appeal of an easy interpretation and human intervention. However, simple network partitions, i.e. between primary distribution (plants to DC's) and secondary distribution (DC's to customers), will require an additional coordinating mechanism.

An alternative approach is based on the empirical computational performance of the models: in an optimal solution, almost all the allocation variables (z_{kb}) take values 0 or 1 naturally, although the location variables (y_k) tend to be fractional. A practical algorithm would be to use an approximate method to assign 0 or 1 values to the location variables (y_k) and drive the allocation variables (z_{kb}) to zero or one by some independent network optimization method. These methods can be combined as Lagrangian heuristic [Cornuéjols, 1991] or cross-decomposition [Van Roy, 1983]: a good solution of a relaxed problem is adapted to the original one.

Among exact or approximate solution methods for CFLP is the relaxation of the capacity constraints (41), yielding an UFLP. In the next part, approximate solutions of large instances of UFLP are proposed, which can help to construct overall solutions of the distribution models.

2. Solution methods for the Uncapacitated Facility Location Problem

Common sense and modelling experience tell us that no exact algorithm will solve all distribution problems efficiently and the theory of computational complexity confirms this intuition. Even decomposing or simplifying the previous examples into common models will not suffice, since the

simplest of the latter is UFLP which is already a NP-hard problem. In oversimplified terms, the time required by existing solution methods grows exponentially with the size of some UFL problem. Indeed, there is no known algorithm which can solve all instances of the simple plant location problem in a time polynomially related to the size of the data (for our purposes the number of potential sites). The conjecture is that no such algorithm can be devised; otherwise, it could be used to solve a large number of reputedly difficult optimization problems.

Given this difficulty, are fast approximations the best recourse left to managers? Actually, a flurry of heuristics has been proposed in the last two decades. Since UFLP is used as a starting board for larger problems, it is useful to review heuristics previously proposed to solve UFLP and compare them with exact algorithms.

2.1 Heuristic algorithms

The initial procedures designed to obtain some solutions to the uncapacitated facility location problem were based on simple heuristic ideas. First and foremost, the bump and shift routine [Kuehn, 1963]: in the add routine, locations are selected one at a time, on the basis of the minimum incremental cost incurred; subsequently, in the bump routine, each location is temporarily removed to see if any savings can be achieved; finally, in the shift routine, each location not selected can be interchanged with one selected if this lowers the total distribution cost. Kuehn and Hamburger's heuristic was seminal in generating a large number of variations, extensions and computational tests: [Manne, 1964; Feldman, 1966; Levy, 1966; Teitz, 1968; Jarvinen, 1972; Diehr, 1972; Hansen, 1972; Eilon, 1978]. Among them [Cornuejols, 1977] branded the procedures greedy (interchange) heuristics. A dynamic programming heuristic incrementing the number of selected sites was proposed in [Baker, 1974]. The natural idea of aggregating the customers in groups and finding a best facility for each group, represented by [Cooper, 1963, 1964; Maranzana, 1964], had less success. Among other heuristics introduced are those of [Drysdale, 1969, Rosing, 1978; Hochbaum, 1982]. All the preceding heuristics can construct a solution from scratch, whether independently or as an initial guess pending refinement by other heuristics or an optimal algorithm.

2.2 Performance of heuristics versus exact algorithms

The failure of heuristics to yield and guarantee an optimal solution is often dismissed by the remark: "Who cares to know the exact solution when the data is so approximate?". [Geoffrion, 1976] partially refutes this tenet by pointing out that most managerial decisions entail the comparison of various alternatives evaluated in several computations:

"Not only does an optimizing capability enhance the value of most individual runs, but it also provides the opportunity to make valid comparisons between the results of different runs. This is extremely important because the conclusions reached by the planning project typically rely far more heavily on comparisons between computer runs than on runs considered individually. With "quasi-optimizing" programs, such as so-called cost calculators or simulators fitted with heuristics, one never knows whether different results are due to different inputs or to the vagaries of the computer program."

For example, Table 1 of Section 2.3 presents a heuristic solution and an optimal value for a problem with 33 potential sites and various levels of fixed charges f listed in the table. For $f=1500$ and 1000 , the heuristic solutions are different, whereas the optimal locations remain unchanged. Faced with an increase of fixed cost, the decision-makers might have relocated their facilities because they relied on a simple heuristic! The value of a heuristic is, therefore, greatly enhanced by a measure of its quality. In fact, for the p -facility location problem, [Cornuejols, 1977] shows that the greedy and greedy interchange solutions can at most deviate by 37 percent from the optimum solution and this deviation can be reached in some (worst) cases. However, the worst cases exhibited do not occur for data sets corresponding to UFLP. Therefore, this result only gives a high estimate for the worst case of UFLP; a case with a 27% deviation can be displayed.

The results of Tables 1, 2 and 3 also indicate that the typical performance of the greedy heuristic is much better than that of the worst-case examples mentioned. A common inference is that worst-cases

Table 1:
Comparative Results for Some 33 x 33 Problems

Fixed Charge	Greedy Heuristic			Erlenkotter's algorithm		
	Objective Value	Number of Locations Chosen	CPU Time (ms)	Optimal Value	Number of Locations Chosen	CPU Time (ms)
5000	27474	2	6	27474	2	7
4000	25474	2	9	25474	2	7
3000	23474	2	6	23474	2	8
2500	22474	2	6	22127	3	9
2000	21166	4	10	20363	4	46
1500	18391	5	11	17832	6	8
1000	15764	6	13	14832	6	4
500	11734	10	22	11267	10	3
295	8710	18	47	8673	17	3
184	6024	31	86	6024	31	4

are contrived and never occur in practice. A different viewpoint is that the problems confronting a given problem-solver form but a small subset of all data occurrences for the UFLP. In this perspective, the probabilistic studies performed to qualify the worst-case analysis [Fisher, 1980; Cornuéjols, 1980] show that it is possible to design some good heuristics for "well-rounded cases". Therefore the problem-solvers can hope to devise an efficient method for their subset (for example, when the number of customers or sites is no greater than three [Cho, 1983]). This remark applies equally well to exact algorithms, as reviewed in the next two sections.

2.3 Efficient optimal algorithms

The efficiency of the dual ascent procedure [Erlenkotter, 1978] is illustrated with standard data sets reported in more detail in Chapter 2, featuring as many potential sites as customers (33, 57 and 100). For each set, ten different levels of uniform fixed charges are presented. Tables 1, 2 and 3 [Thizy, 1985] show that the computational time required to solve some problems heuristically or exactly by the dual adjustment procedure, requires approximately similar computational times as the heuristic. More precisely, in 47% of the cases, Erlenkotter's algorithm takes no more time to find the exact

Table 2:
Comparative Results for Some 57 x 57 Problems

Fixed Charge	Greedy Heuristic			Erlenkotter's algorithm		
	Objective Value	Number of Locations Chosen	CPU Time (ms)	Optimal Value	Number of Locations Chosen	CPU Time (ms)
5000	38742	2	13	38547	3	62
4000	36366	3	14	35547	3	38
3000	32559	4	22	32136	4	40
2500	30160	5	19	30022	5	24
2000	27660	5	23	27222	6	23
1500	24700	6	24	23943	7	19
1000	20815	9	50	20307	9	37
500	15348	13	62	15261	13	28
200	9251	29	163	9142	29	7
50	2821	55	393	2821	55	8

Table 3:
Comparative Results for Some 100 x 100 Problems

Fixed Charge	Greedy Heuristic			Erlenkotter's algorithm		
	Objective Value	Number of Locations Chosen	CPU Time (ms)	Optimal Value	Number of Locations Chosen	CPU Time (ms)
8000	90579	4	40	87889	4	226
7000	86579	4	39	83720	5	372
6000	82579	4	40	78720	5	337
5000	78127	5	48	73073	6	273
4000	72599	6	60	66407	7	221
3000	63176	8	94	59407	7	2248
2900	62133	8	99	58613	8	1604
2000	52056	10	113	50103	12	147
1150	39637	14	196	38479	16	141
1000	37435	15	188	35965	17	79

solution than the greedy heuristic. The ratio is even larger - 60% - for the smallest problems (33 clients). The greedy heuristic is considerably faster only in two cases, which can in fact be solved more efficiently by subsequent improvements of Erlenkotter's procedure [Van Roy, 1982; Körkel, 1989]. Other primal-dual methods are explored in [Holmberg, 1995]. Despite their apparent advantage, primal and dual methods can be further enhanced by an extensively streamlined simplex algorithm, following the guidelines set forth by [Cornuejols, 1978; Simao, 1982].

By successfully incorporating subgradient optimization techniques and promoting the advent of dual adjustment procedures, the algorithmic improvement of UFLP has been extremely representative of the main advances of discrete optimization. Its role in polyhedral combinatorics, another largely

successful avenue of optimization, has been more checkered. A partial description of the facets of the simple plant location polytope was undertaken early by [Guignard, 1980; Cornuejols, 1982 (Math. Prog.); Cho, 1983]. Yet, few attempts have been made to design an algorithm that would selectively add constraints based on this description [Aardal, 1994; Van Hoesel, 1994]. Such an algorithm promises to yield important advances for the solution of large locational problems.

For the p -median location problem, it is generally possible to find a (uniform) set of fixed charges that ensures the opening of the desired number p of facilities, and Erlenkotter's algorithm is applicable; this procedure is studied in [Mavrides, 1979] and used in computational experiments [Christofides, 1982]; the method has always been successful with the 3 problem sets reported. Otherwise, [Cornuejols, 1977] presents an efficient subgradient algorithm, [Galvao, 1980] introduces a dual-adjustment procedure and [Christofides, 1982] presents a method outperforming Erlenkotter's algorithm on a graph with 200 nodes.

2.4 The Challenge of Large Uncapacitated Facility Location Problems

In spite of the increasing power of optimal algorithms, heuristics are still needed to solve large instances of UFLP for major strategic problems:

"There exist telecommunication network problems which could use algorithms handling problems with thousands of "plants" and "destinations". These can only be tackled by heuristics at present." [Guignard, 1977]

This is also the problem range for which judicious methods can provide the most dramatic improvements. Of particular interest are the formulations introduced by [Cornuejols, 1980] that solve various linear relaxations with a simplex algorithm, confirming the apparent superiority of dual approaches:

"In all cases dual formulations are much easier to solve than the corresponding primal formulations. In several cases the primal required more than 10 times as many pivots as the dual."

Additional efficiency is often gained by capitalizing on some special structure of the problem. For example, the data sets of Tables 1, 2 and 3 possess a Euclidian distance matrix, for which the following characteristics are not exploited by Erlenkotter's algorithm:

symmetry: the distance from City i to City j is equal to the distance from City j to City i . City i can be serviced from a site in City i at no cost.

triangularity: the sum of the distances from City i to City j and from City j to City k is not smaller than the distance from City i to City k .

These features are typically present in the p -median problem. Other particularities are the density of the underlying transportation network (i.e., whether all customers can be serviced from all potential sites, or on the contrary, only a few routes are permitted), and the implicit treatment of unequal demands (via a multiplication of the corresponding unit costs of distribution), practically favouring potential sites close to large demand areas.

Contrary to the greedy heuristic, the length of the exact resolution is not directly related to the number of facilities selected: [Khumawala, 1972] reports that the problem seems easier when only a few sites are chosen or rejected. In the latter case, the Stingy Heuristic proposed in [Feldman, 1966], which opens all locations initially and closes them one at a time if this proves economical, is faster than its greedy counterpart.

Yet, the capacity to exploit some special structure of the problem restricts the applicability of exact algorithms for large problem sizes. Consider for example an occurrence of CFLP with symmetric costs. Lagrangian relaxation to reduce it to UFLP destroys the symmetry in the objective function (9.w) and renders a customized algorithm inapplicable. Given the current state-of-the-art, exact algorithms lose their edge over heuristics for general data sets as occur in the distribution problems

presented. Table 3 points to two cases of such weakness, where the time required by the exact algorithm increases rapidly and the greedy heuristic is considerably faster. The heuristic actually yields an approximate solution in a time polynomially bounded by the problem size.

Chapter 2

Analysis of a simple heuristic for the uncapacitated facility location problem

In spite of the progress of exact algorithms to solve large instances of the uncapacitated facility location problem, fast approximate solutions serve decision-makers well in many cases delineated below.

- The data is so approximate that a rough-cut solution will suffice. Even in this case, a heuristic will be useful only if no comparison is required with alternative scenarios, as pointed out at the beginning of Section 2.2 of the previous chapter.
- An evaluation of various scenarios is needed in an environment so uncertain that it is better assessed by a large series of quick, approximate but direct measures of alternatives than an encompassing sensitivity analysis.
- The size of the problem is beyond the reach of optimal algorithms and decision-makers must reconcile themselves to the preceding two remedies.

However, with the loss of accuracy, decision-makers need to know how much faith they can put in the approximation. Therefore, the value of a heuristic is greatly reinforced by a measure of its quality. Measurable approximation of discrete optimization problems is a fairly new area of research because it cannot rely on the classical results that apply to continuous functions. Hence, the interest for the approximate solution of such optimization problems has spurred the design of a large number of algorithms, some of which yield a solution with a guaranteed degree of accuracy. In particular, the greedy solution to the uncapacitated facility location problem is shown [Cornuéjols, 1977] never to deviate from the optimal solution by more than $\frac{1}{e} \approx .37$ (as a relative error), where e is the base of

natural logarithms. This result has drawn much attention for its conceptual simplicity and the relative quality of the approximation. In this chapter, the results of [Cornuéjols, 1977] are compared with those of [Cornuéjols, 81], using data sets reported in [Thizy, 1997].

1. The greedy heuristic for the uncapacitated facility location problem

The formulation of the K-facility location problem, a variant of the uncapacitated facility location and the p-median problems, is recapitulated below, with $m = |I|$ and $n = |J|$. Since the cost coefficients can be positive or not, a maximization objective is chosen to conform with a large part of the literature on the heuristic.

$$z = \max \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}$$

$$\sum_{j=1}^n x_{ij} = 1 \quad \text{for } i = 1, \dots, m \quad (1)$$

$$\sum_{j=1}^n y_j \leq K \quad (2)$$

$$0 \leq x_{ij} \leq y_j \leq 1 \quad \text{for } i = 1, \dots, m \quad (3)$$

In Constraint (2), the parameter p used in Chapter 1 is replaced by the letter K to clearly distinguish it from an index and to produce formulas that can be easily compared with classical results.

The greedy heuristic yields a feasible solution of the problem by choosing facilities one at a time; at each step, the chosen facility is one that produces the largest improvement in the objective value per unit of investment, without considering potential synergies with or among locations by-passed. To formalize this principle, the following algorithm is given:

The greedy location algorithm [Kuehn, 1963]

k counts the number of locations chosen (recorded in J^*)

Step 1 Let $k = 1$, $J^* = \emptyset$ and $u_i^1 = \min_{j=1, \dots, n} c_{ij}$ for $i = 1, \dots, m$

Step 2 Let $\rho_j(u^k) = \sum_{i=1}^m \max(0, c_{ij} - u_i^k)$, $j \in J^*$

Find $j_k \in J^*$ such that $\rho_{j_k}(u^k) = \max_{j \in J^*} \rho_j(u^k)$.

If $|J^*| > K$ or $\rho_{j_k}(u^k) \leq 0$ and $|J^*| \geq 1$ go to Step 4.

Otherwise set $J^* = J^* \cup \{j_k\}$ and go to Step 3.

Step 3 Set $k = k + 1$. For $i = 1, \dots, m$, set $u_i^k = \max_{j \in J^*} c_{ij}$

$$= u_i^{k-1} + \max(0, c_{ij_{k-1}} - u_i^{k-1})$$

Go to Step 2.

Step 4 Stop; set $H = k-1$, the greedy solution is given by $y_j = 1$, $j \in J^*$ and $y_j = 0$ otherwise, and the value of the greedy solution is

$$z_G = \sum_{i=1}^m u_i^1 + \sum_{h=1}^H \rho_{j_h}$$

In this equation and in the sequel of the chapter, ρ_{j_h} denotes $\rho_{j_h}(u^h)$. Actually, for further notational

simplicity, ρ_h will be used instead of $\rho_{j_h} [= \rho_{j_h}(u^h)]$. Let $z_R = \sum_{i=1}^M u_i^1$: this value represents a lower

estimate of the worst possible choice of location, i.e., a selection with not even a simple calculation.

It serves as a benchmark for the performance of the heuristic which performs myopic calculations.

Property 1 [Cornuéjols, 1977]:

For the K-median location problem, the greedy heuristic satisfies

$$\frac{z - z_G}{z - z_R} \leq \left(\frac{K-1}{K} \right)^K \left(< \frac{1}{e} \right)$$

The following section introduces a new bound on the accuracy of the greedy heuristic for the K-median problem, which can be shown to be asymptotically equivalent to the preceding bound.

2. A new bound for the greedy location heuristic

Property 2 [Cornuéjols, 1981]: Consider the simple K-plant location problem (1) - (3) with matrix $C=(c_{ij})$ of zero-one costs. Let d be the maximum number of non-zero entries in a column of C and k be the smallest number such that:

$$\frac{1}{d} + \frac{1}{d-1} + \dots + \frac{1}{k+1} \leq 1$$

then the value z_G of the greedy solution is related to the optimal value z of the problem by

$$\frac{z - z_G}{z} \leq \frac{k}{d} \left(\frac{1}{d} + \dots + \frac{1}{k+1} \right)$$

A tighter bound than that of Property 2 holds. However, the improvement is only known after the greedy solution has been calculated.

Corollary 1 [Cornuéjols, 1981]: Consider Problem (1) - (3) with zero-one costs c_{ij} .

$$\frac{z - z_G}{z} \leq \frac{h}{d} \left(\frac{1}{d} + \frac{1}{d-1} + \dots + \frac{1}{h+1} \right)$$

where $h = \rho_k$ when the greedy heuristic terminates. The applicability of Property 2 is heightened by an alternative formulation of the uncapacitated facility problem. Both the derivation and the ensuing numerical results resemble the canonical representation of UFLP [Cornuéjols, 1980, 1984].

3. A disaggregation of the uncapacitated facility location problem

For a given set of potential sites, replace the set I of clients (whose demand must be satisfied) by a set of profit margins, as shown in the following example using five sites and one client. Let the profits generated by satisfying the demands from the five locations be:

$$1 \quad 3 \quad -2 \quad 0 \quad 1$$

Decompose the profits into two incremental levels:

$$\begin{array}{ccccc} 1 & 2 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{array}$$

Repeat the procedure:

$$\begin{array}{ccccc} 1 & 1 & -2 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}$$

Successive disaggregations yield the final reduction:

$$\begin{array}{ccccc} -2 & -2 & -2 & -2 & -2 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}$$

The uppermost row will be accounted in the reference value z_R . The procedure can be applied to a second, third client, etc. A description of the procedure for cost matrix C is summarized below:

a) Row reduction

Step 0. If all the elements c_{ij} of Row i equal 0, delete the row and stop. Else go to Step 1.

Step 1. Let r_i be the second largest element in Row i . Add a new row (c_{i1}, \dots, c_{in}) to the matrix, where $c_{ij} = 1$ if $c_{ij} > r_i$, 0 otherwise. Subtract this new row from Row i . Go to Step 0.

b) Matrix reduction

Perform row reduction for all $i = 1, \dots, m$.

One can show [Cornuéjols, 1978, 1980, 1984] that each solution $(x_j)_{j \in J}$ of the linear relaxation of (P) is also a solution of the linear relaxation of the UFLP with the new set of clients.

Using the preceding reduction, Property 2 can be extended to the general problem (1) - (3) where the coefficients of the objective function can have any integral value since the greedy heuristic on the associated zero-one canonical problem chooses the same locations.

Corollary 2 [Cornuéjols, 1981]: Consider Problem (1) - (3). Let $d = \max_{j=1, \dots, n} \sum_i c_{ij}$ and k the

smallest integer such that:

$$\frac{1}{d-z_R} + \frac{1}{d-1-z_R} + \dots + \frac{1}{k+1} \leq 1.$$

Then

$$\frac{z-z_G}{z-z_R} \leq \frac{k}{d-z_R} \left(\frac{1}{d-z_R} + \frac{1}{d-z_R-1} + \dots + \frac{1}{k+1} \right).$$

Corollary 2 can be restated in terms of the original problem (1) - (3):

Corollary 3 [Cornuéjols, 1981]:

Consider Problem (1) - (3). Let k be the largest integer such that:

$$1 - \frac{\rho_2}{\rho_1} + 1 - \frac{\rho_3}{\rho_2} + \dots + 1 - \frac{\rho_k}{\rho_{k-1}} \leq 1;$$

then the value z_G of the greedy solution is related to the optimal value z of the problem by

$$\frac{z - z_G}{z - z_R} \leq \frac{\rho_k}{\rho_1} \left(1 - \frac{\rho_2}{\rho_1} + \dots + 1 - \frac{\rho_k}{\rho_{k-1}} \right)$$

Note again that this bound can only be calculated after the heuristic has been performed, unlike the bound of Property 2, obtained a priori. Section 4 will show that the additional information gained during the computation does sharpen the measure of accuracy of the greedy heuristic.

4. Computational results

Computational results are presented in order to assess whether Corollary 2 yields a practical advance over Property 1. First, the bound proposed Property 1 is calculated in Table 1 for various values of K . Note the convergence toward $1/e$. Table 1 shows that the bound is tighter for small values of K (of course, for $K=1$, the greedy heuristic yields an optimal solution). Hence, the subsequent analysis will focus on small values of K .

K	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$(K-1/K)^K$	0	25	30	32	33	33	34	34	35	35	35	35	35	35	36

Table 1: Percentage values of the bound of Property 1

The bound is now compared numerically with the bound cited in Corollary 2, using calculations and comments reported in [Thizy, 1997]. The data tests follow a standard suite for K-facility location problems patterned after [Ahn, 1998] and also found in [Simao, 1989]. They feature as many potential sites as customers: 33 and 57 major U.S. cities in two sets of problems initially proposed by [Karg, 1964], and 100 points randomly selected in a rectangle for the third set. These problems, solved in [Schrage, 1975; Cornuéjols, 1977; Erlenkotter, 1978; Cornuéjols, 1982 (SIAM)], are deemed to be good representatives of K-facility location problems. Table 2 presents results consistent with Table 1. The refinement of the bound is significant for small values of K, and seems to be enhanced by larger problem sizes. More extensive evidence of this performance is given by a large set of randomly generated problems. The first set of results involves random points in 10 squares of each size: 10^i for $i=1, 2, 3$ and 4 , each square containing 70 or 100 points connected by integer Euclidian distances. Each row of Table 3 reports results averaged over 10 data sets generated randomly. Table 3 confirms the experimental results of [Ahn, 1988; Simao, 1989] that found the data sets showed little sensitivity to the size of the squares, i.e, the rounding of the data. One set of problems with only 70 nodes indicates that the range of accuracy does not vary widely with the number of nodes. This behaviour will be set against further results reported in subsequent tables .The second set of results is based on random trees generated with $n = 50, 100$ or 150 nodes, using the method described in [Even, 1979]. All edge lengths are set equal to 1 in the trees. Each row of Table 4 also cites results averaged over 10 data sets generated randomly. In this case, the accuracy appears to improve with the number of nodes, again in accordance with the experimental results of [Ahn, 1988; Simao, 1989] that found larger problems to display better linear relaxations.

$n \backslash K$	1	2	3	4	5	6	7	8	9	10	11	12
30	0	16	25	20	21	28	28	28	28	29	30	30
33	0	20	22	25	28	28	31	32	32	32	32	32
57	0	11	23	24	24	31	31	31	31	31	31	31
100	0	14	18	22	22	22	26	32	32	33	32	33

Table 2: Percentage values of the bound of Corollary 2 for standard data sets.

n	$\bar{n}K$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
70		0	12	17	27	29	27	28	29	30	30	31	31	32	32	32
100	1	0	12	18	27	28	27	27	29	29	30	31	31	32	32	32
100	2	0	11	14	27	29	29	28	29	29	30	30	32	32	32	33
100	3	0	11	15	27	29	28	28	28	29	30	31	32	32	33	33
100	4	0	11	15	27	29	28	28	28	29	30	31	32	32	33	33
100	5	0	11	15	27	29	28	28	28	29	30	31	32	32	33	33

Table 3: Percentage Values of the bound of Corollary 2 for complete Euclidian graphs

Another set of results is based on trees which do not have edges of length 1; their nodes are constituted by 30 to 200 random points in the unit square, using the scheme explained earlier. The length of an edge in the tree is the Euclidean distance between its endpoints rounded to an integer. The distance between two nodes of the tree is the length of the unique path joining them. The results contained in Table 4 are confirmed over a larger range of problem sizes that test the influence of their size on the accuracy of the estimate. Each row of Table 5 still reports results averaged over 10 data sets generated randomly. The results resemble those of Table 4, confirming that the accuracy does not appear to be influenced by the type of distance function, as found by the experimental results of [Ahn, 1988; Simao, 1989].

$n \setminus K$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
50	0	20	21	21	21	24	26	26	28	30	30	30	30	30	30
100	0	17	19	18	22	23	25	26	27	28	29	30	31	31	32
150	0	13	17	20	23	26	25	26	27	28	29	29	30	30	30

Table 4: Percentage values of the bound of Corollary 2 for trees with unit costs

$n \backslash K$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
30	0	20	22	24	25	27	28	28	29	30	31	31	31	31	32
50	0	20	22	24	26	28	29	31	29	29	29	30	30	31	31
70	0	18	22	24	24	26	28	29	30	31	31	32	32	32	32
100	0	18	20	23	27	27	27	29	28	28	28	29	30	31	32
150	0	15	23	24	25	28	28	29	29	31	31	31	31	31	32
200	0	13	20	24	25	28	28	28	29	31	31	31	32	32	32

Table 5: Percentage values of the bound of Corollary 2 for Euclidian trees

The third set of results pertains to sparse random graphs. The first results correspond to lengths equal to 1, where a sequence of graphs is generated, starting from a random tree with 50 nodes, then adding at each i -th iteration 50 random edges to the graph of the previous iteration. The entries of Table 6 entail one such graph and various values of i . Each row of Table 6 presents results for only one tree with 50 nodes. Table 6 indicates that the accuracy of the estimate decreases at first with the density of the tree, then tapers off and may possibly improve with even larger density. Again, the results echo the increased difficulty of problems, measured by duality gap, as a function of tree density. Even the moderate reversal of the effect is noted in the results of [Ahn, 1988; Simao, 1989].

The results further extend to sparse random graphs in the unit square. The cost between two adjacent nodes is the Euclidian distance, and the shortest path between non-adjacent nodes. As in Table 6, each row of Table 7 involves an additional set of n arcs. The accuracy has increased for $k = 2$ with

n	$n \backslash K$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
50	0	0	20	22	23	26	29	18	18	19	19	22	24	24	28	32
50	1	0	16	23	23	24	25	29	29	29	29	29	29	31	31	31
50	2	0	19	20	20	23	28	32	32	32	33	33	33	33	33	33
50	3	0	18	25	25	28	29	29	31	33	34	34	34	34	34	34
50	4	0	18	21	21	28	31	32	32	33	33	33	33	33	33	33

Table 6: Percentage values of the bound of Corollary 2 for sparse graphs with unit cost

n	$\bar{n}K$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
50	1	0	11	25	28	29	31	31	32	32	33	33	33	34	34	34
50	2	0	8	25	27	29	29	29	31	32	32	32	32	32	32	32
50	3	0	8	17	27	27	27	27	27	30	30	33	33	34	34	34
50	4	0	10	20	28	28	28	30	30	32	32	32	32	33	33	33
50	5	0	10	22	27	27	27	29	32	32	32	32	32	33	33	33
50	6	0	10	20	28	29	29	29	30	33	33	33	33	34	34	33
50	7	0	11	24	26	26	27	27	27	30	30	30	30	33	33	33
50	8	0	10	23	29	31	30	31	31	31	32	32	32	33	33	33
50	9	0	11	23	29	30	28	28	28	28	28	28	28	28	32	32
50	10	0	11	25	28	29	28	28	28	28	29	29	29	29	33	33
50	11	0	12	24	29	30	28	28	28	28	29	29	29	33	33	33
50	12	0	11	20	29	29	26	26	27	27	29	29	29	33	33	33
50	15	0	12	17	28	31	26	27	27	27	28	28	28	32	31	31
50	17	0	12	17	26	31	29	28	29	29	29	29	31	31	31	31
100	0	0	12	13	23	29	29	28	29	30	30	30	32	33	33	33
100	5	0	13	19	27	28	27	31	32	33	33	33	33	33	33	34
100	10	0	11	19	26	25	25	25	25	26	31	32	33	33	33	33
100	20	0	13	13	26	26	26	29	30	30	30	26	26	26	26	27
100	25	0	15	25	26	29	31	31	31	32	31	31	31	31	32	31
100	28	0	15	24	26	28	31	32	32	31	31	31	32	32	32	32
100	30	0	15	24	26	28	31	31	31	31	31	31	32	32	32	32
100	33	0	15	24	25	25	25	28	28	31	31	31	31	32	33	32
100	40	0	15	25	25	28	30	30	32	33	33	33	33	33	33	33
100	42	0	15	25	25	28	30	30	32	33	33	33	33	33	33	33
150	0	0	25	27	23	22	22	22	28	29	32	33	32	32	32	32
150	10	0	14	25	29	23	23	23	24	24	29	29	31	33	33	34
150	20	0	13	24	29	29	29	30	30	31	31	31	32	32	32	32
150	30	0	14	24	25	27	28	30	30	30	30	30	30	31	31	33
150	40	0	18	21	21	22	30	31	31	31	32	32	32	33	33	33
150	45	0	14	24	28	30	30	26	26	27	27	27	27	27	27	31

Table 7: Percentage values of the bound of Corollary 2 for sparse Euclidian graphs

respect to the tree of Table 6. To confirm the results, and assess the influence of the number of nodes, larger trees ($n=100$ and 150) are examined. Although the data sets reveal little sensitivity to the size of the squares, i.e, the rounding of the data, the accuracy is still better than that of Table 6, not only for $k=2$ but for slightly larger values as well. The effect of density is apparently slighter for larger trees, a behaviour not tested in the experiments of [Ahn, 1988; Simao, 1989].

Finally, the fourth set of experimental results deals with the uniform cost model, based on a set of 30 problems with random integer costs. For $i=1, 2$ and 3 , several sets S_i of 10 graphs with 100 nodes feature costs uniformly distributed in the range $[1, 10^i]$. Each row of Table 7 contains set averages. The results are in contrast with those of the preceding tables, as the accuracy has markedly diminished for small values of k . On the other hand, as in Table 3, the data sets show little sensitivity to the size of the squares, i.e, the rounding of the data, as noted in the experimental results of [Ahn, 1988; Simao, 1989]. Also, as in Table 3, the range of accuracy does not vary widely with the number of nodes, an effect not tested in the experiments of [Ahn, 1988; Simao, 1989].

A global view of the results shows that the improvement brought by the new bound of Property 2 is particularly important for small values of k , the number of facilities selected. Therefore, Property 2 ensures decision-makers that the heuristic solution is closer to the optimum than could be guaranteed before.

n	$\backslash K$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
70		0	21	24	25	28	29	30	30	30	31	31	31	31	30	29
100	1	0	21	25	26	28	29	30	31	31	31	31	31	31	31	31
100	2	0	21	25	26	28	29	30	30	31	31	32	32	32	32	32
100	3	0	21	25	26	28	29	30	30	31	31	32	32	32	32	32
200		0	21	25	26	29	29	30	31	31	31	32	32	33	33	33

Table 8: Percentage values of the bound of Corollary 2 for random costs

Chapter 3

Aggregate and average network capacity management

The previous chapter presented the difficulty of approximating even the simplest version of a distribution model. A guaranteed measure of accuracy of the heuristic eventually rested on a network disaggregation. In fact, network capacity aggregation can also yield valid approximations while reducing the size of the optimization problem. In this chapter, the combinatorial difficulty caused by an NP-hard problem such as UFLP is replaced by a functional one: the optimal solution of the aggregate problem does not have a simple expression and must be approximated. This chapter also relies on a simple example, this time taken from the specialized logistic field of inventory control. A section is devoted to the well-known EOQ model. Another section, addressing stochastic elements, attempts to estimate the approximate effect that some change in the level of demand has on the optimum cost of a continuous review (Q,r) inventory control system.

1. Homogeneous capacity planning

In many distribution problems, processing facilities have uniform capacities, i.e.:

- a) location of standard facilities such as concentrators [Mirzaian, 1985],
- b) fleet of vehicles [Magnanti, 1981],
- c) production planning [Krarup and Bilde, 1977; Yano, 1985 (U. Mich)].

Case a) can be modelled with a formulation of the capacitated facility location problem, equivalent to the formulation introduced in Chapter 1 and stated below.

$$z = \min \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + \sum_{j \in J} f_j y_j, \quad (1)$$

$$\sum_{j \in J} x_{ij} = 1 \quad \text{all } i \in I, \quad (2)$$

$$\sum_{i \in I} x_{ij} \leq s_j y_j \quad \text{all } j \in J, \quad (3)$$

$$y_j = 0 \text{ or } 1 \quad \text{all } j \in J. \quad (4)$$

Suppose that some facilities $j \in J_0 \subset J$ have equal characteristics; accordingly, drop the index j from the fixed costs f and capacities s . The problem can be relaxed by aggregating Expressions (1) and (3) for $j \in J_0$, which yields:

$$z = \min \sum_{i \in I} \sum_{j \in J_0} c_{ij} x_{ij} + f \sum_{j \in J_0} y_j + \sum_{i \in I} \sum_{j \in J \setminus J_0} c_{ij} x_{ij} + \sum_{j \in J \setminus J_0} f_j y_j \quad (5)$$

$$\sum_{i \in I} \sum_{j \in J_0} x_{ij} \leq s \sum_{j \in J_0} y_j \quad (6)$$

Let $y = \sum_{j \in J_0} y_j$. Expressions (4), (5) and (6) can be re-written:

$$z = \min \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} + f y + \sum_{i \in I} \sum_{j \in J \setminus J_0} c_{ij} x_{ij} + \sum_{j \in J \setminus J_0} f_j y_j$$

$$\sum_{j \in J} \sum_{i \in I} x_{ij} \leq s y \quad \text{all } j \in J,$$

$$y \geq 0 \text{ integer}$$

Such a problem occurs either naturally because of physical and economic characteristics, i.e. with off-the-shelf telephone switches, a pool of interchangeable trucks or in capacity planning with uniform capacity increments. There are several interpretations of the preceding formulation. In network design, suppose that a facility such as represented in Figure 1 can be segmented. The facility can now be viewed as the aggregation of homogeneous modular facilities, as represented in Figure 2. Since in Figure 2, each homogeneous facility j represents a segment of the central facility, one can assume

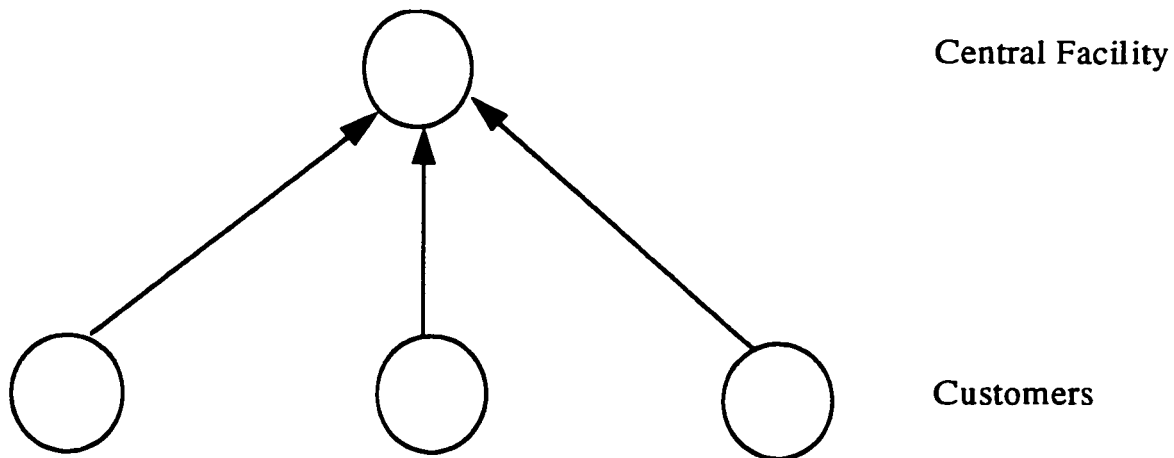


Figure 1: Location of a central facility

that the processing variable costs c_{ij} do not vary between any facility $j \in J_0$; denote them c_i . Expression (5) becomes:

$$z = \min \sum_{i \in I} c_i \sum_{j \in J_0} x_{ij} + fy + \sum_{i \in I} \sum_{j \in \mathcal{N}J_0} c_{ij} x_{ij} + \sum_{j \in \mathcal{N}J_0} f_j y_j$$

Denote also $x_i = \sum_{j \in J_0} x_{ij}$. The program (1)-(4) becomes:

$$z = \min \sum_{i \in I} c_i x_i + fy + \sum_{i \in I} \sum_{j \in \mathcal{N}J_0} c_{ij} x_{ij} + \sum_{j \in \mathcal{N}J_0} f_j y_j \quad (7)$$

$$x_i + \sum_{j \in \mathcal{N}J_0} x_{ij} = 1 \quad \text{all } i \in I,$$

$$\sum_{i \in I} x_{ij} \leq s_j y_j \quad \text{all } j \in \mathcal{N}J_0,$$

$$\sum_{i \in I} x_i \leq sy, \quad (8)$$

$$y_j = 0 \text{ or } 1$$

$$\text{all } j \in J \setminus J_0,$$

$$y \geq 0 \text{ integer .}$$

(9)

The remainder of the chapter attempts to treat the elements of the problem solely related to the aggregate facility, by focusing on the constraints of the previous formulation where these elements can be easily isolated, i.e. (7), (8) and (9). The resulting system is further simplified to:

$$z = \min cx + fy,$$

(10)

$$x \leq sy$$

$$y \geq 0 \text{ integer}$$

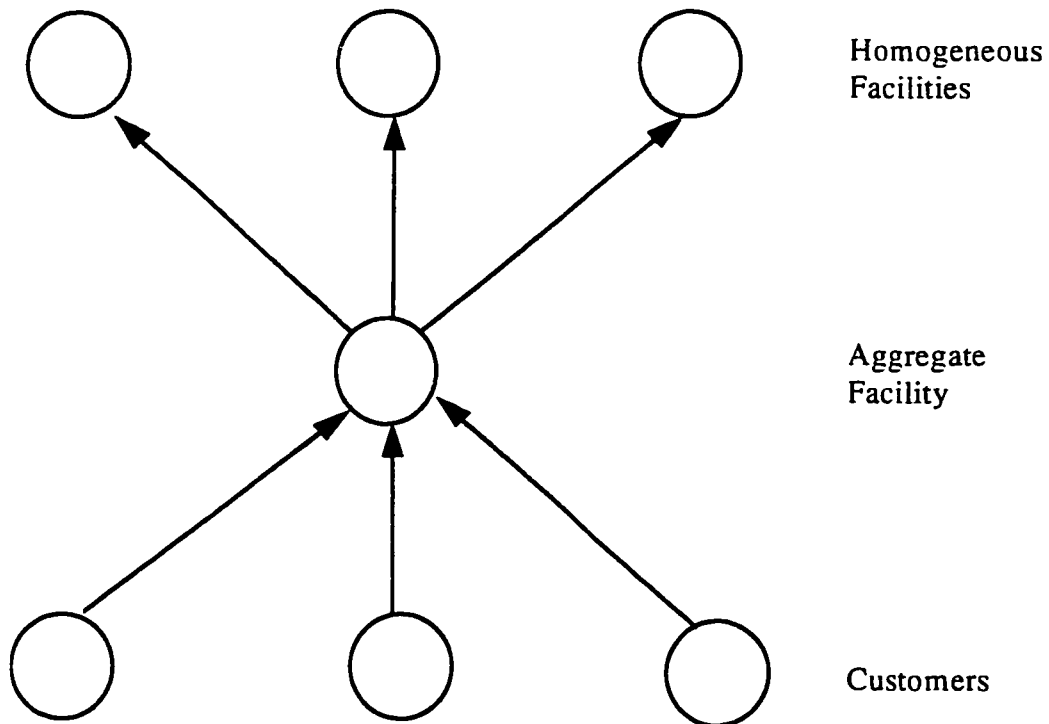


Figure 2: Location of homogeneous facilities

Each segment of the central facility can also be viewed as a capacity increment. A different representation of such increments is common in capacity planning. The cost of establishing a processing centre is depicted in Figure 3, in which:

- s is the maximum capacity of each increment,
- f is the fixed cost of establishing each increment (e.g start-up, set-up costs, lump payment),
- the slope of the graph along each segment indicates the variable processing cost incurred along each increment.

An additional question of capacity planning is to determine the size of the increment s , which becomes a variable to which a cost H can be assigned. This addition makes the problem markedly more difficult to solve. However, it is possible to give a simple approximation that uses the effect of (dis)aggregation. With a cost minimization criterion, each increment j will be used to capacity s before a fixed cost f is incurred to use Increment $j+1$. If there are many increments, the cost function can be asymptotically approximated by a linear function. In particular, assume that the loss incurred by not using the last increment j fully can be ignored, hence Equation (8) is approximated as $x = ys$.

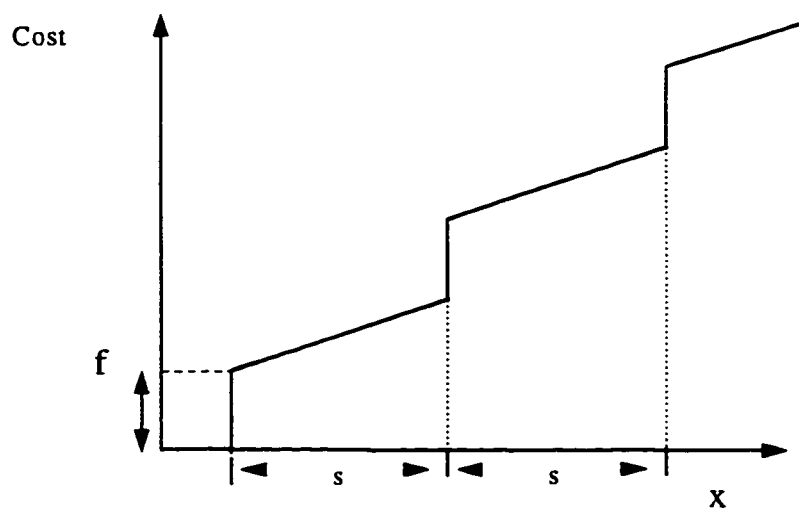


Figure 3: Repetitive fixed cost function

This additional equality can be used to express one variable as a function of the other three. The objective function (10) becomes:

$$z = \min cx + Hs + fx/s$$

which will be abbreviated as:

$$z = \min \mathcal{F}(x, s)$$

The penalty for this simplification is the introduction of a non-linear objective function. In many distribution problems, as in this simple case, the presence of incremental capacities is fairly independent of the rest of the problem and can be simplified in this way. Further, if the capacity of the aggregate facility does not appear in additional constraints, it may be omitted as an explicit parameter of the objective function by treating its own cost as a function of x , thus yielding the more compact expression:

$$z = \min F(x)$$

Of course, similar disaggregations can occur for other central facilities. Exemplifying the approach by a classical model of inventory control, the following sections display properties of the optimal solution of the problem and approximate its expression when the demand is stochastic. Note a change of notation to keep within the tradition of literature on inventory control models:

- c will be renamed T ,
- H will be replaced by $h/2$,
- f will be renamed K ,
- s will be renamed Q ,
- x will be considered as a parameter D that will be later allowed to vary.

A brief, but comprehensive survey of fixed order quantity policy models is given first.

2. Repetitive fixed charge problem: batch manufacturing

Operations management distinguishes two general modes of physical processing:

- process flows such as chemical treatment, paper or food manufacturing, in which materials undergo a continuous transformation;
- batch processing where materials are temporarily collected as lots; the lot is then treated in one batch.

Such lots create stocks of material awaiting processing. Disregarding the characteristics of the production of the batch itself, the section simplifies the problem of intermittent processing by emphasizing the management of such recurring stocks.

2.1 Review of Inventory Control

Although the term inventory originally refers to a physical verification of the level of stock held at a given date, it is commonly accepted in the following definition. An inventory is a collection of items or resources which can be stored for future use. Typical contents of inventory consist generally of raw materials, finished products, component parts, supplies or work in process. From a practical standpoint, keeping inventories has several purposes:

- to maintain independence of operations,
- to absorb variation in product demand,
- to allow flexibility in production scheduling,
- to provide a safeguard against variation in raw material delivery time and
- to take advantage of economic purchase order size.

An inventory policy is a set of rules to determine which items should be maintained in inventory and in what quantities, when stock should be replenished, how large and frequent orders should be. An

inventory system comprises actions and controls that implement the inventory policies. Many types of inventory problems arise in logistic management. They can be grouped in broad classes according to the number of different product types, the pattern of demand (deterministic, varying in time, etc.), the number of planning periods and their repetitiveness, the nature of lead time and replenishment modes, the allowance for stockouts and recourse to back ordering. Most inventory policies are based on reducing, or even minimizing inventory costs, which are traditionally decomposed as:

- holding (carrying) costs that comprise maintenance, cost of space and storage equipment, spoilage, loss, cost of capital, labour and management,
- setup (production change) costs, ordering costs, and
- shortage costs, including loss of goodwill.

The last item is difficult to assess and the next section shows how this measurement is often circumvented. Among many inventory policies, the following sections focus on continuous review, fixed order quantity models that determine at which dates orders of pre-specified size are placed. A continuous review inventory control system is a process in which an order of fixed quantity Q is placed as soon as the inventory position, given by the inventory on hand plus the inventory on order less the back orders, drops to a fixed reorder point r . When the demand process is given as a random variable, the stochastic model is referred to as the (Q,r) model.

2.2 A Fixed Order Quantity Model for Deterministic Demand

As a simple benchmark, a classical model generally known as the economic order quantity (EOQ) [Harris, 1915; Erlenkotter, 1989, 1990] is first reviewed and later compared to a stochastic version.

The model is based on a set of very restrictive assumptions:

- the demand is constant, uniform and deterministic,
- replenishment lead time is constant, deterministic,
- the unit product price is constant and deterministic,

- the ordering cost is constant,
- the ordered quantity is delivered in a unique shipment and
- no back orders are allowed.

The variations of inventory level are depicted in Figure 4, displaying two parameters:

- τ , the lead time and
- r , the reorder point,

Other parameters of the model are:

- K , the ordering (setup) cost,
- h , the holding cost per unit of time,
- C , the cost per unit
- D , the demand per unit of time,
- Q , the quantity to be ordered.

With this notation, the total cost is made up of three components:

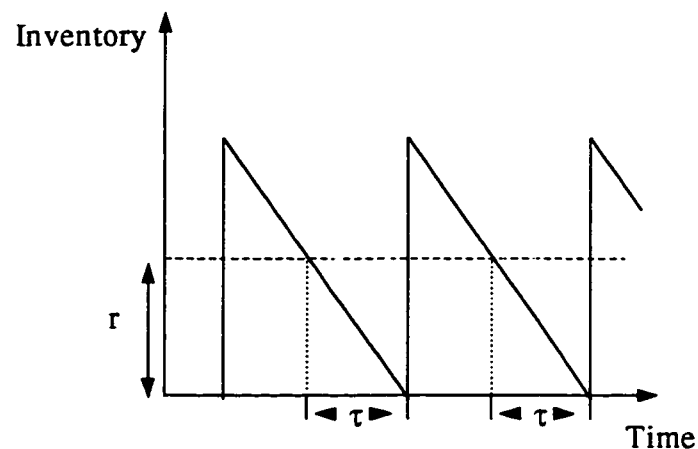


Figure 4: A Fixed order model for constant demand

- DC , the annual purchase cost,
- $(D/Q)K$, the annual cost of making (D/Q) orders,
- $(Q/2)h$, the annual holding cost given an average inventory level of $Q/2$ units.

Hence, the overall inventory costs T_c can then be expressed as:

$$T_c = DC + \frac{D}{Q} K + \frac{Q}{2} h$$

The cost components and the total cost T_c are represented in Figure 5 as a function of the order size Q to be determined. The optimal values Q_{opt} and T_c^* can be derived via a simple differentiation:

$$\frac{\partial T_c}{\partial Q} = \frac{\partial(DC + \frac{D}{Q}K + \frac{Q}{2}h)}{\partial Q} = 0$$

$$\rightarrow -\frac{DK}{Q^2} + \frac{h}{2} = 0$$

$$\rightarrow Q^2 = \frac{2DK}{h}$$

$$\rightarrow Q_{opt} = \sqrt{\frac{2DK}{h}}$$

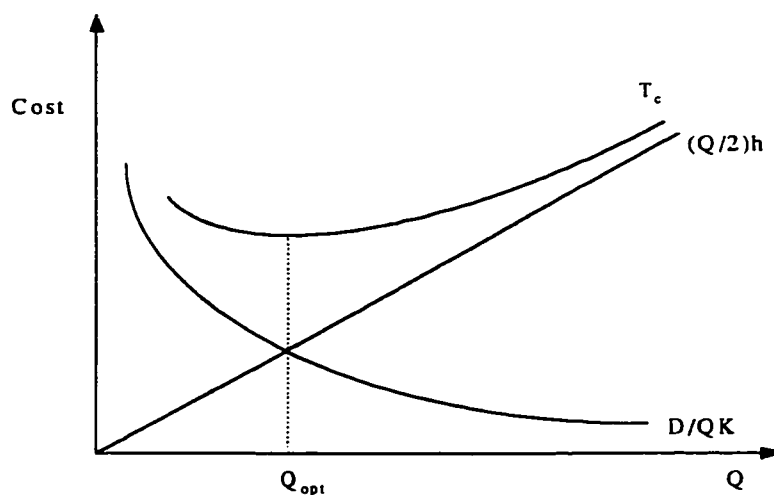


Figure 5: Cost Components of the EOQ Model

$$T_C^* = DC + \sqrt{2DKh}$$

and the optimal number of replenishments per year can be expressed as:

$$N_{opt} = \frac{D}{Q_{opt}} = \frac{D}{\sqrt{\frac{2DK}{h}}} = \sqrt{\frac{Dh}{2K}}$$

In addition to the original goal of describing a closed-form solution of the problem, it is worth showing empirically that the expression of Q_{opt} is robust, i.e., yields a cost function T_C that varies little as the values of the variables or the parameters change. Consider a base case with $D = 10,000$ units; $K = \$10$ per order; $h = \$0.20$. The optimal order quantity is:

$$Q_{opt} = \sqrt{\frac{2(10000)10}{0.2}} = 1000 \text{ units}$$

Table 1 below shows how the size of the quantity affects each type of inventory cost. Equally significant are the variations of total costs with changes in parameters of the model. Next is an example displaying this influence. If the actual ordering cost was: $K = \$20$ per order, the optimal order quantity would be:

$$Q_{opt} = \sqrt{\frac{2(10000)20}{0.2}} = 1414.21 \text{ units}$$

and the corresponding total cost:

$$T_C^* = \frac{10000}{1414} (20) + \frac{1414}{2} (0.2) \approx \$282.80$$

Q	Ordering cost	Holding cost	Total cost
962	104.0	96.2	200.2
977	102.4	97.7	200.1
996	100.4	99.6	200.0
997	100.3	99.7	200.0
998	100.2	99.8	200.0
999	100.1	99.9	200.0
1000	100.0	100.0	200.0
1001	99.9	100.1	200.0
1002	99.8	100.2	200.0
1003	99.7	100.3	200.0
1004	99.6	100.4	200.0
1023	97.8	102.3	200.1
1040	96.2	104.0	200.2

Table 1: Variation of Inventory Cost with Order Quantity

Given that 1000 units were to be ordered, the total cost would be:

$$T_c^* = \frac{10000}{1414} (20) + \frac{1414}{2} (0.2) \approx \$282.80$$

Hence, the relative difference of total cost is only:

$$\frac{T_c - T_c^*}{T_c^*} = \frac{300 - 282.80}{282.80} = 0.061$$

Thus, a 100% variation of K resulted only in a 6.1% increase in total costs. However, more complex models of inventory control must be devised for uncertain demand.

3. Distribution network with stochastic demand

In traditional logistic management, models as those of Chapter 1 consider average network characteristics such as demands, because they support strategic planning, whereas inventory control is classified as an operational exercise to face stochastic elements. Even though cost averaging will somewhat respect the stochastic nature of demand, capacity averaging is typically insufficient and must be accompanied by additional safety buffers. This problem is reviewed in the next section.

The original goal of the analysis carried out in the preceding section was to simplify the objective function within the more general framework of capacity planning:

$$z = \min cx + Hs + fx/s$$

by expressing z as a function of x :

$$z = \min F(x) = cx + 2 \sqrt{Hfx} .$$

In the preceding section, the expression of the optimal value $F(x)$ of the inventory policy resulted from unconstrained optimization; this expression was shown to be robust. In the sequel, more detail is required, for example when the variables x and s can be involved in additional constraints, preventing such a compact representation. In the following section, the effect of a simple constraint is analysed; the example of stochastic inventory control is chosen not only for its mathematical properties, but for an illustration of a treatment of stochastic events in a logistic application.

Stochastic programming methods have been applied to the design of logistic networks under fluctuating demand [Mulvey, 1989, 1991]. Within the focus of the thesis, i.e., assessing the potential contribution of well-known models, we consider again the standard techniques used for inventory control, as they have a double relevance to logistic management under stochastic demand, offering:

- a general approach to repetitive fixed charge problems as occur in capacity planning and

- an immediate interest to logistic design which must address feasibility and cost issues raised by recurring operations.

3.1 Fixed Order Quantity Models for Stochastic Demand

This section describes a fixed-order quantity inventory system, designed to satisfy a random demand. Since the exact amount of demand is unknown, a safety stock must be maintained to control the occurrence of shortages, as shown in Figure 6. A safety stock is an inventory that is carried to ensure that the random fluctuations in demand do not force a stockout in a period where immediate ordering cannot replenish the inventory. The size of the safety stock will depend on how responsive the firm wishes to be, on the demand variability and on the tradeoff between holding costs and shortage costs. Classical analyses of inventory policies aim at minimizing the (long-run) average total costs (the order cost, inventory holding cost and the back order penalty cost) per unit time. In practice, the determination of the shortage cost is considered to be very difficult [Oral, 1972]. Therefore the service level is generally measured either via an estimated probability of (avoiding) stockout or an expected amount of stockout. To illustrate the difference between the two criteria, suppose that each month, the safety stock maintained at ordering time at a warehouse is exactly 4. Table 2 contains a random sample of demands during lead-times and resulting shortages.

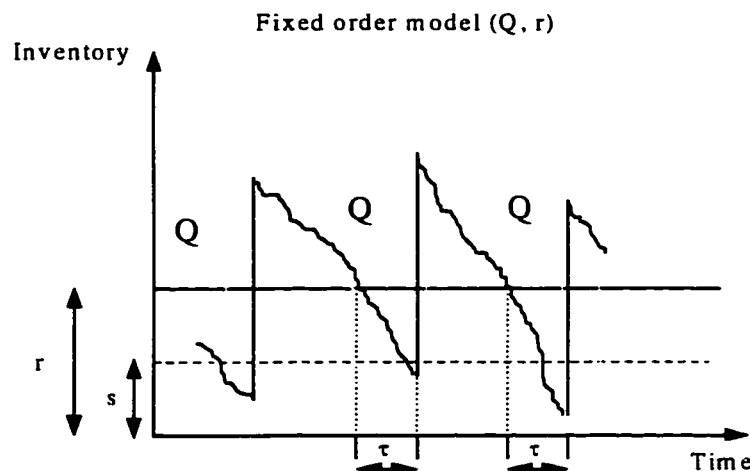


Figure 6: Variation of Inventory in a (Q,r) System

Lead-Time Quantity	
Requested	Short
5	1
7	3
5	1
1	0
2	0
7	3
3	0
1	0
2	0
5	1
3	0
2	0

Table 2: Random sample of shortages

One can classify the shortages according to their magnitude and list the frequency of each occurrence, yielding Table 3 that summarizes Table 2.

Lead-Time Quantity		Relative Frequency
Requested	Short	
7	3	2/12
5	1	3/12
3	0	2/12
2	0	3/12
1	0	2/12

Table 3: Tally of order sizes

3.1.1 Service level based on frequency

A first measure of service level is based on the probability P that inventory will be sufficient to meet the demand D_τ occurring during the replenishment lead time τ . From the sample, the estimated probability of avoiding stockout is $P = 7/12$, as shown in Table 4. Conversely, to reach a predetermined service level characterized by P , a reorder point r can be calculated via $P(D_\tau \leq r) = P$.

Service level based on frequency for a standard normal distribution

A typical assumption is that the rate of demand follows a normal distribution with standard deviation σ_D . Then the demand during lead time follows a normal distribution with the following parameters:

$$D_\tau \sim N(D\tau, \sigma_\tau^2 = \tau\sigma_D^2)$$

Reducing the latter to a standard normal distribution, the service level can be characterized by:

$$z = \frac{r - D\tau}{\sigma_D \sqrt{\tau}}$$

Lead-Time Quantity		Annual Frequency	Frequency of satisfactory service
Requested	Short		
7	3	2/12	0
5	1	3/12	0
3	0	2/12	2/12
2	0	3/12	3/12
1	0	2/12	<u>2/12</u>
			7/12

Table 4: Cumulative frequency of stock availability

i.e., a number of acceptable normal standard deviations. Using a variable Z with a standard normal distribution, the value of z can be found via: $P(Z \leq z) = P$, e.g. in a standard normal table. Reciprocally, the reorder level r required to achieve the service level characterized by z (indexing P) is:

$$r = D\tau + z\sigma_D \sqrt{\tau}$$

3.1.2 Service level based on expected unsatisfied demand

Unfortunately, a measure of service level based on the probability of avoiding stockout during the replenishment lead time τ , does not recognize the fact that stockouts may have varying magnitudes, as shown in Table 5. In answer, the measure adopted in the sequel is based on the average number of units that cannot be supplied from stock currently on hand, e.g. 9/12 in Table 5.

Following a classical definition of a (Q,r) policy, the service level s is based on the expected unsatisfied demand that will occur during the replenishment lead time τ (i.e. demand above the level of safety stock). Define:

	Lead-Time Quantity		Frequency	Expected Shortage
	Requested	Short		
7	3	2/12	6/12	
5	1	3/12	3/12	
3	0	2/12	0	
2	0	3/12	0	
1	0	2/12	<u>0</u>	
			9/12	

Table 5: Expected shortage

$$1 - s = \frac{b(r)}{Q}$$

where:

$1-s$ is the percentage of unsatisfied demand, ($0 \leq s \leq 1$).

$b(r) = \int_{x=r}^{\infty} (x - r) f(x)dx$ is the expected lead-time demand above the reorder point and

$f(x)$ is the probability density function for the lead-time demand.

The service level is therefore a joint function of Q and r . The established practice is to calculate Q as if demand was deterministic, i.e., as Q_{opt} , according to the EOQ formula.

Service level for a standard normal distribution

Applying the preceding assumptions to a normally distributed rate of demand D , the measure of service level now specializes to:

$$1 - s = \frac{E(z)\sigma_D}{Q_{opt}} \sqrt{\tau}$$

where $E(z) = \int_{x=z}^{\infty} (x - z) \phi(x)dx$ and $\phi(x)$ is probability density function of the standard normal

distribution. Reciprocally, the level of service can be indexed by z calculated from the expression:

$$E(z) = \frac{(1 - s)Q_{opt}}{\sigma_D \sqrt{\tau}}$$

In practice, the value z is directly found from $E(z)$ in a table established for a standard normal

distribution [Brown, 1967]. Then, the reorder point r is determined via z by:

$$r = D\tau + z\sigma_D\sqrt{\tau}$$

Example:

Cost to place an order	$K = \$30$
Holding Cost	$h = \$60/\text{unit}/\text{year}$
Annual demand	$D = 500 \text{ units} / (365 \text{ days}/\text{year})$
Standard deviation	$\sigma_D = 2 \text{ unit}/\text{day}$
Lead time:	$\tau = 25 \text{ days}$
Service level	$s = 97\%$

$$Q = \sqrt{\frac{2DK}{h}} = \sqrt{\frac{2 \times 500 \times 30}{60}} = 22.36$$

$$E(z) = \frac{(1-s)Q}{\sigma_D\sqrt{\tau}} = \frac{(1-.97)22.36}{10} = .066,$$

From which one can obtain the numerical approximation: $z = 1.12$. Hence:

$$r = D\tau + z\sigma_D\sqrt{\tau} = \frac{500}{365} \times 7 + 1.12 \times 10 = 20.79$$

3.2 Sensitivity Analysis of Continuous Inventory Control with Service Level Constraint

This section analyses an approximation of the continuous review inventory control system. The rate of demand need not follow a normal distribution, but it is assumed, as in typical (Q,r) inventory

control systems, that unsatisfied orders are entirely back ordered [Hadley and Whitin, 1963]. The objective is also to minimize the (long-run) average total cost per unit time while respecting a pre-specified service level.

The inventory control problem can be formulated as a program minimizing the approximate sum of order and holding costs, subject to a service level constraint. Although the problem can be optimized, analytical expressions of the optimal values Q and r are unavailable. On the other hand, it is possible to give a closed form expression of the derivatives of the optimum value. This will enable a sensitivity analysis with respect to the level of demand D , with the goal of expressing the optimal value as a function of D . Many analyses on optimal ordering policies select some special form of demand pattern [Schneider, 1978; Bachmann, 1984; Yano, 1985 (NRLQ)]. The sensitivity analysis of [Zheng, 1992] uses back order costs rather than a service level.

3.2.1 Mathematical programming formulation

The optimization problem of determining an optimal (Q,r) policy can be stated as follows [Ng, 1997]:

$$F(Q,r) = \min_{Q,r} \frac{KD}{Q} + h \left[\frac{Q}{2} + r - \mu \right] \quad (11)$$

such that

$$\frac{b(r)}{Q} \leq (1 - s) \quad (12)$$

$$Q > 0 \quad (13)$$

where μ is the expected demand during lead time τ : $\mu = \int_{x=0}^{\infty} x f(x) dx$.

In the above formulation, Constraint (13) restricts the model to policies maintaining a positive reorder quantity; Constraint (12) specifies a service level expected by requiring that the shortage during the lead time should not exceed a percentage $(1-s)$ of the order.

The expression of the objective cost function and the service level constraint are based on standard approximations that ignore the small gain in inventory cost produced by stockouts [Hadley, 1963; Johnson, 1974]. A similar system is proposed in [Aardal, 1989], with the additional constraint $\mu - r < 0$. This assumption is restrictive and omitted from the model analysed here.

Assume that the probability density function for the lead-time demand $f(x)$ does not depend on Q ; therefore $b(r)$ will not depend on Q . Properties 4 and 5 below will assume that it does not depend on other parameters of the models: K, h, s . The next lemma shows that $b(r)$ is a convex function of r .

Lemma [Ng, 1997]:

$$\frac{\partial^2 b(r)}{\partial r^2} \geq 0.$$

Proof: Denote by $\Phi(x)$ the cumulative distribution function for the lead-time demand; by

definition $b(r) = \int_{x=r}^{\infty} (x-r) f(x) dx = [(x-r)(\Phi(x) - 1)]_r^{\infty} - \int_{x=r}^{\infty} [\Phi(x) - 1] dx = \int_{x=r}^{\infty} [1 - \Phi(x)] dx$, hence

$$\frac{\partial b(r)}{\partial r} = \Phi(r) - 1 \text{ and } \frac{\partial^2 b(r)}{\partial r^2} = f(r) \geq 0.$$

Constraint (12) is re-written as:

$$b(r) - Q(1 - s) \leq 0.$$

First, recall the Kuhn-Tucker conditions for a convex program. Let x^* be a relative minimum point for the problem:

$$\begin{array}{ll} \text{minimize} & F(x) \\ \text{subject to} & g(x) \leq 0 \end{array}$$

Then there is a vector $\lambda \geq 0$ such that:

$$\begin{aligned} \nabla F(x^*) + \lambda^T \nabla g(x^*) &= 0 \\ \lambda^T g(x^*) &= 0 \end{aligned}$$

Denote $g_1(r, Q) = b(r) - Q(1 - s)$ and $g_2(r, Q) = -Q$; the system of Kuhn-Tucker conditions contains the following equalities:

$$\frac{\partial F}{\partial Q} + \lambda \frac{\partial g}{\partial Q} = 0$$

$$\frac{\partial F}{\partial r} + \lambda \frac{\partial g}{\partial r} = 0$$

$$\lambda g(Q, r) = 0$$

$$\lambda \geq 0$$

3.2.2 Sensitivity analysis

The preceding expressions can serve as a base to calculate the effect that some change in problem parameters has on the optimum solution (Q,r) and its value $F(Q,r)$. In [Ng, 1997] the preceding system of Kuhn-Tucker equations is differentiated with respect to the general parameter u , as:

$$\left[\frac{\partial^2 F}{\partial Q^2} + \lambda \frac{\partial^2 g}{\partial Q^2} \right] dQ + \left[\frac{\partial^2 F}{\partial Q \partial r} + \lambda \frac{\partial^2 g}{\partial Q \partial r} \right] dr + \frac{\partial g}{\partial Q} d\lambda + \left[\frac{\partial^2 F}{\partial Q \partial u} + \lambda \frac{\partial^2 g}{\partial Q \partial u} \right] du = 0$$

$$\left[\frac{\partial^2 F}{\partial Q \partial r} + \lambda \frac{\partial^2 g}{\partial Q \partial r} \right] dQ + \left[\frac{\partial^2 F}{\partial r^2} + \lambda \frac{\partial^2 g}{\partial r^2} \right] dr + \frac{\partial g}{\partial r} d\lambda + \left[\frac{\partial^2 F}{\partial r \partial u} + \lambda \frac{\partial^2 g}{\partial r \partial u} \right] du = 0$$

$$\lambda \frac{\partial g}{\partial Q} dQ + \lambda \frac{\partial g}{\partial r} dr + g(Q,r) d\lambda + \lambda \frac{\partial g}{\partial u} du = 0$$

Intuitively, any positive increase in the parameters K , h , s , τ or D of the problem should increase the optimal value of the objective function. Increases in K and decreases in h should intuitively increase the optimal value of Q ; for such variations, the value of b is proportional to that of Q and varies in the opposite direction to r . Thus, increases in h and decreases in K should intuitively increase the optimal value of r . Increases in D should also increase the optimal values of Q and r . Finally, increases in s and τ are expected to increase the optimal value of r . This intuition is confirmed by the analytical solutions obtained by solving the preceding system of equations, and substituting u by specific parameters of the optimization program, as shown in the following properties [Ng, 1997].

Property 1. The (largest) optimal value r is an non-increasing function of K , while that of Q and the objective value are non-decreasing functions of K .

Property 2. The optimal values of the objective value and the (largest) value of r is a non-decreasing function of h , while that of Q is a non-increasing function of h .

Property 3. The (largest) optimal values of r and the objective value are non-decreasing functions of s .

Property 4. The (largest) optimal values Q , r and the objective value are non-decreasing functions of τ .

A last property addresses the influence of a variation of demand on the optimal values of the program. Property 5 will differ from the preceding four because the variation in the rate of demand affects simultaneously two parameters of the original program: D and $b(r)$. To analyse the effect of a variation of expected shortage during the lead-time $b(r)$, the notation needs to be slightly refined. In keeping with the assumption of constant average demand, assume now that demand is likely to occur at equal rates at all times of the order cycle (including lead time). Therefore the distribution function for the lead-time demand is $f_{\delta\tau}(x) = f_1(x/\delta\tau)/\delta\tau$, where δ characterizes the magnitude of the demand rate and f_1 , a reduced distribution function, depends on none of the other parameters of the model. Hence $\Phi_{\delta\tau}(r) = \Phi_1(r/\delta\tau)$, $b_{\delta\tau}(r) = \delta\tau b_1(r/\delta\tau)$, $\mu = \delta\tau\mu_1$ and $\frac{\partial b_{\delta\tau}(r)}{\partial r} = \frac{\partial b_1(r_1)}{\partial r_1}$, where the reduced

expressions, $D_1 = D_{\delta\tau}/\delta$, $\Phi_1(x) = \int_{t=-\infty}^x f_1(t)dt$, $r_1 = r/\delta\tau$, $b_1(p) = \int_{x=p}^{\infty} (x-p)f_1(x)dx$ and $\mu_1 = \int_{t=0}^{\infty} tf_1(t)dt$ depend

on none of the other parameters of the model. Hence, Property 5 considers the combined influence of δ upon $b_{\delta\tau}(r)$ and $D_{\delta\tau}$.

Property 5. The optimal value Q and the objective value are non-decreasing functions of δ . The (largest) optimal value r is a non-decreasing function of δ for $r \geq -b(0)$ and non-increasing otherwise.

The variation of the objective function (11) can actually be given an analytic expression as a function of δ . Its derivation is based on a customization of the general method [Ng, 1997] that yielded the five preceding properties. The preceding Kuhn-Tucker system can be written in detail as:

$$-\frac{KD}{Q^2} + \frac{h}{2} - \lambda_1 (1-s) - \lambda_2 = 0 \quad (14)$$

$$h + \lambda_1 \frac{\partial b(r)}{\partial r} = 0 \quad (15)$$

$$b(r) - Q(1-s) \leq 0 \quad (16)$$

$$Q > 0 \quad (17)$$

$$\lambda_1 [b(r) - Q(1-s)] = 0 \quad (18)$$

$$\lambda_2 Q = 0 \quad (19)$$

$$\lambda_2 \geq 0 \quad (20)$$

$$\lambda_1 \geq 0 \quad (21)$$

Since $b(r) \geq 0$ and $s \leq 1$, Constraint (16) enforces $Q \geq 0$, and Constraint (14) forces $Q > 0$. Thus, $\lambda_2 = 0$ from Constraint (19), which simplifies the system of constraints (14)-(21) by eliminating Constraints (17), (19) - (20) and simplifying Constraint (14) to:

$$-\frac{KD}{Q^2} + \frac{h}{2} - \lambda_1 (1-s) = 0 \quad (22)$$

From Constraint (15), $\frac{\partial b(r)}{\partial r} < 0$, and therefore $\lambda_1 \geq 0$ holds without having to impose Constraint (21)

explicitly. By eliminating the formally unconstrained variable λ_1 , the system (15),(22) becomes:

$$\frac{\partial b(r)}{\partial r} \left(\frac{KD}{h Q^2} - \frac{1}{2} \right) = 1-s \quad (23)$$

The resulting Kuhn-Tucker system is formed by Constraints (16), (23) and:

$$h[b(r) - Q(1-s)] = 0 \quad (24)$$

Constraint (24) can then simply be expressed as $b(r)=(1-s)Q$, which makes Constraint (16) redundant and also helps to eliminate Q from Equation (23), thus yielding as Kuhn-Tucker conditions:

$$\frac{\partial b(r)}{\partial r} \left(\frac{KD(1-s)}{h b^2(r)} - \frac{1}{2(1-s)} \right) = 1$$

or:

$$\frac{\partial b(r)}{\partial r} - \left(\frac{KD(1-s)}{h b^2(r)} - \frac{1}{2(1-s)} \right)^{-1} = 0 \quad (25)$$

for a pre-specified function $b(r)$. Although Equation (25) does not provide any analytic expression of r in the general case, it will serve to assess the effect of variation in the parameters of the problems, such as demand or lead time.

First, differentiating Equation (25) with respect to a general parameter u , yields:

$$U = \frac{\partial^2 b(r)}{\partial u \partial r} + \left(\frac{\partial b(r)}{\partial r} \right)^2 \left[\frac{(1-s)}{b^2(r)} \frac{\partial}{\partial u} \left(\frac{KD}{h} \right) - \frac{2KD(1-s)}{h b^3(r)} \frac{\partial b(r)}{\partial u} - \left(\frac{KD}{h b^2(r)} + \frac{1}{2(1-s)^2} \right) \frac{\partial s}{\partial u} \right] \quad (26)$$

in particular, when $u = r$, Expression (23) simplifies to:

$$R = \frac{\partial^2 b(r)}{\partial r^2} - \frac{2KD(1-s)}{h} \left(\frac{1}{b(r)} \frac{\partial b(r)}{\partial r} \right)^3.$$

With these expressions, the variation of optimal reorder level is: $dr/du = -U/R$.

To specialize the expression (23) of U when $u = \delta$, first note:

$$\frac{\partial^2 b(r)}{\partial r^2} = \frac{1}{\delta \tau} \frac{\partial^2 b_1(r_1)}{\partial r_1^2},$$

$$\frac{\partial b(r)}{\partial \delta} = \left[b_1(r_1) - r_1 \frac{\partial b_1(r_1)}{\partial r_1} \right] \frac{\partial}{\partial \delta} (\delta \tau) = \left[b(r) - r \frac{\partial b(r)}{\partial r} \right] \frac{1}{\delta \tau} \frac{\partial}{\partial \delta} (\delta \tau) = \left[b(r) - r \frac{\partial b(r)}{\partial r} \right] \frac{1}{\delta}, \text{ and}$$

$$\frac{\partial^2 b(r)}{\partial r \partial \delta} = -\frac{r_1}{\delta \tau} \frac{\partial^2 b_1(r_1)}{\partial r_1^2} \frac{\partial}{\partial \delta} (\delta \tau) = -\frac{r}{\delta \tau} \frac{\partial^2 b(r)}{\partial r^2} \frac{\partial}{\partial \delta} (\delta \tau) = -\frac{r}{\delta} \frac{\partial^2 b(r)}{\partial r^2} .$$

Therefore Expression (26) becomes:

$$U = \frac{1}{\delta} \left\{ -r \frac{\partial^2 b(r)}{\partial r^2} + \left(\frac{\partial b(r)}{\partial r} \right)^2 \left[KD \frac{(1-s)}{hb^2(r)} - 2KD \frac{(1-s)}{hb^3(r)} \left(b(r) - r \frac{\partial b(r)}{\partial r} \right) \right] \right\}$$

$$U = -\frac{1}{\delta} \left[r \frac{\partial^2 b(r)}{\partial r^2} + \left(\frac{\partial b(r)}{\partial r} \right)^2 KD \frac{(1-s)}{hb^2(r)} \left(1 - \frac{2r}{b(r)} \frac{\partial b(r)}{\partial r} \right) \right]. \quad (27)$$

From Equation (24), the variation of Q and r as a function of an independent parameter u is regulated

$$\text{by the equation: } (1-s) dQ = \frac{\partial b(r)}{\partial r} dr + \frac{\partial b(r)}{\partial \delta} d\delta. \quad (28)$$

Recall:

$$F(Q,r) = \min_{Q,r} \frac{KD}{Q} + h \left(\frac{Q}{2} + r - \mu \right)$$

and therefore:

$$dF(Q,r,\delta) = \left(\frac{h}{2} - \frac{KD}{Q^2} \right) dQ + h dr + \left(\frac{K}{Q} \frac{\partial D}{\partial \delta} - h \frac{\partial \mu}{\partial \delta} \right) d\delta \quad (29)$$

Applying Equations (27) and (28) to (29) yields:

$$(1-s) dF(Q,r,\delta) = \left(\frac{h}{2} - \frac{KD}{Q^2} \right) \left(\frac{\partial b(r)}{\partial r} dr + \frac{\partial b(r)}{\partial \delta} d\delta \right) + h(1-s) dr + (1-s) \left(\frac{K}{Q} \frac{\partial D}{\partial \delta} - h \frac{\partial \mu}{\partial \delta} \right) d\delta$$

which, by collecting terms, becomes:

$$(1-s) dF(Q,r,\delta) = \left[\left(\frac{h}{2} - \frac{KD}{Q^2} \right) \frac{\partial b(r)}{\partial r} + h(1-s) \right] dr + \left[\left(\frac{h}{2} - \frac{KD}{Q^2} \right) \frac{\partial b(r)}{\partial \delta} + (1-s) \left(\frac{K}{Q} \frac{\partial D}{\partial \delta} - h \frac{\partial \mu}{\partial \delta} \right) \right] d\delta$$

or, noting that the first term cancels:

$$\begin{aligned} \frac{dF(Q,r,\delta)}{d\delta} &= -h \frac{\partial r}{\partial b(r)} \frac{\partial b(r)}{\partial \delta} + \frac{K}{Q} \frac{\partial D}{\partial \delta} - h \frac{\partial \mu}{\partial \delta} \\ &= \frac{K \partial D}{Q \partial \delta} - h \left(\frac{\partial r}{\partial b(r)} \frac{\partial b(r)}{\partial \delta} + \frac{\partial \mu}{\partial \delta} \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\delta} \left\{ \frac{KD}{Q} - h \left[\frac{\partial r}{\partial b(r)} \left(b(r) - r \frac{\partial b(r)}{\partial r} \right) + b(0) \right] \right\} \\
&= \frac{1}{\delta} \left\{ \frac{KD}{Q} - h \left[\frac{\partial r}{\partial b(r)} b(r) - r + b(0) \right] \right\} \\
&= \frac{h}{\delta} \left[\frac{b(r)}{2(1-s)} + r - b(0) \right] \\
&= \frac{h}{\delta} \left[\frac{Q}{2} + \int_{-\infty}^r \Phi(x) dx \right] \geq 0 \tag{30}
\end{aligned}$$

With the presence of a service level constraint, the previous expression (30) is not simple enough to yield a direct expression of $F(Q,r, \delta)$ itself. In particular, as in the unconstrained case, the optimal cost $F(Q,r, \delta)$ is not a linear expression of the demand.

Returning to the aggregate logistic planning of Chapters 1 and 2, Expression (30) seems to invalidate the fairly standard assumption that the operational costs of distribution are proportional to the amount of flow, e.g., the demand at the warehouse. Indeed, storage operations entailing exactly one product would allow economies of scale consistent with Expression (30), practically ushering the ideal target of "streamlined inventory management." However, typical inventory management involves a variety of products and even if only one physical product was handled, the various modes of operations (automatic vs. manual control of bill-of-lading, location of vehicle unloading, position of shelves, times of inspection, etc.), would fragment this common product into a variety of subfamilies, each governed analytically by an independent system (11)-(13). An increase in volume is generally

accompanied by new versions, sizes or grades which increase the number of related systems (11)-(13) and effectively render inventory costs somewhat proportional to the flow of products distributed. Naturally, as seen in Equations (16) of Chapter 1, bounds can be placed on operations so as to validate linear approximations.

Conclusion

A contribution of this thesis is to bring together two approaches to analyse fairly different types of approximations arising in logistic planning. First, it was shown how the greedy heuristic yielded a better degree of accuracy than was known previously for problems involving a choice of a few locations, thereby alleviating the combinatorial complexity of discrete location. Then, following a direction usually reserved to the operational level of logistic planning, simple solutions of a more detailed capacity planning problem were first observed in the traditional EOQ model of inventory control and subsequently in its stochastic extension. In spite of a series of nested approximations, the presence of a constraint introduced a functional complexity which is partially workable.

Modelling and solution issues

Some logistic problems require an optimal closure of sites or, when re-designing a distribution network, most potential sites may need to be selected as locations of facilities. In this restrictive case, a greedy heuristic should yield a good guarantee of accuracy since it may only discard a few sites erroneously. Yet, the measure of accuracy presented in Chapter 2 did not yield any noticeable improvement over the classical result of $(1/e)$. A possible avenue is to study a drop heuristic such as the stingy heuristic that discards iteratively the least promising location. Since the stingy heuristic performs well in practice, but the only known bound on its accuracy is .5, an ex-post measure may improve the theoretical measure considerably.

Property 5 of Chapter 3 presented an expression of inventory cost as a function of the level of demand. The function is not linear, whereas strategic models of distribution, as presented in Chapter 3, assume a linear relationship. It seems natural to introduce non-linear terms in the objective function of such models. On the other hand, these expressions themselves are approximate. Therefore, piecewise linear approximations may suffice, especially as the range of variation of operations is limited. Widely different ranges are often viewed as operations of different facilities (e.g. local versus regional warehousing, even if they are physically connected).

Directions of Future Research

By referring to a hierarchical approach to the planning of logistic operations, the thesis opens a number of issues mostly found in the area of hierarchical production planning, some of which delineated below.

- What level of aggregation should be adopted for a first-cut, strategic approach to the selection of major features of the operations?
- How much leeway should be given to more detailed logistic models, e.g., those that model stochastic aspects of operations?
- Is pricing (via Lagrangian multipliers) sufficient to ensure some coordinating mechanism between the levels of planning or, on the other hand, must constraints such as capacity levels reflect the higher levels of decision?

The thesis also reveals the relative value of some of the terms of each model of optimization, and therefore the difficulty of coordinating each level of a logistic model. For example, at the detailed level of capacity planning and inventory control, the difficulty of finding optimal solutions is caused by the presence of fixed costs; the same terms can be aggregated as variable costs at a higher level. Conversely, some fixed costs of a strategic level are simply removed from the operational models. Fortunately, the very difficulty of accommodating all these terms in an all-encompassing model enhances the benefit of a hierarchical approach.

Bibliography

Aardal, K., "Polyhedral results for the uncapacitated facility location problem: lifting and separation," *Mathematical Programming Symposium*, Ann Arbor, 1994.

Aardal, K., O. Jonsson and H. Jonsson, "Optimal inventory policies with service level constraints," *Journal of the Operational Research Society*, 40, 1989, 65-73.

Aardal, K., Y. Pochet and L.A. Wolsey, "Capacitated Facility Location: Valid Inequalities and Facets," *Mathematics of Operations Research*, 20(3), 1996, 562-582.

Ahn, S., C. Cooper, G. Cornuejols and A. Frieze, "Probabilistic analysis of a relaxation for the K-median problem," *Mathematics of Operations Research*, 13(1), 1988, 1-31.

Aikens, C.H., "The Optimal Design of a Physical Distribution System on a Multicommodity Multi-Echelon Network," The University of Tennessee, Knoxville, Tennessee, 1982.

Aikens, C.H., "Facility Location Models for Distribution Planning," *European Journal of Operational Research*, 22, 1985, 263-279.

Albright, S.C., "An approximation to the stationary distribution of a multiechelon repairable-item inventory system with finite sources and repair channels," *Naval Research Logistics Quarterly*, 36(2), 1989, 179-195.

Aly, A.A. and B. Rahali, "Analysis of a bicriteria location model," *Naval Research Logistics Quarterly*, 37 (6), 1990, 937-944.

Arcelus, F.J., "On the selection of industry location alternatives," *Eng. Costs and Prod. Econ*, 16(3), 1989, 213-219.

Bachmann, A., "Implementing a periodic review (r,Q) model with constraints," *Proceedings of the Third International Symposium on Inventories*, Budapest, 1984.

Baker, K.R., "A Heuristic Approach to Locating a Fixed Number of Facilities," *Logistics and Transportation Review*, Autumn 1974.

Balinski, M.L. and K. Spielberg, "Methods for Integer Programming: Algebraic, Combinatorial, and Enumerative," in Aronofsky, J.S. (editor): *Progress in Operations Research III*, John Wiley and Sons, New-York, 1969, 195-292.

Balinski, M.L. and P. Wolfe, "On Benders Decomposition and a Plant Location Problem," Working Paper ARO-27, *Mathematica*, Princeton, New Jersey, 1963.

Beasley, J.E., "A Note on Solving Large p-Median Problems," Technical Report, Department of Management Science, Imperial College, London, England, September 1984.

Beasley, J., "An algorithm for solving large capacitated warehouse location problems," *European Journal of Operational Research*, Netherlands, 33(3), 1988, 314-325.

Beckmann, M.J., C.B. Mcguire and C.B. Winsten, "Studies in the economics of transportation," *Yale University Press*, New Haven, CT, 1956.

Bhaskaran, S. and M.A. Turnquist, "Multiobjective transportation considerations in multiple facility location," *Transportation Research*, 24A(2), 1990, 139-148.

Bilde, O. and J. Krarup, "Sharp Lower Bounds and Efficient Algorithms for the Simple Plant Location Problem," *Annals of Discrete Mathematics*, 1, 1977, 79-97.

Boffey, T.B., "Location problems arising in computer networks," *Journal of the Operational Research Society*, 40(4), 1989, 347-354.

Brandeau, M.L. and Chiu, S.S., "An overview of representative problems in location research," *Management Science*, 15(6), 1989, 645-674.

Brown, R.G., *Decision rules for inventory management*, Holt, Rinehart & Winston, New-York, 1967.

Captivo, M.E., "Fast primal and dual heuristics for the p-median location problem," *European Journal of Operational Research*, The Netherlands, 52(1), 1991, 65-74.

Cho, D.C., E.L. Johnson, M.W. Padberg and M.R. Rao, "On the Uncapacitated Plant Location Problem I: Valid Inequalities and Facets," *Mathematics of Operations Research*, 8(4), 1983, 579-589.

Cho, D.C., M.W. Padberg and M.R. Rao, "On the Uncapacitated Plant Location Problem II: Facets and Lifting Theorems," *Mathematics of Operations Research*, 8(4), 1983, 590-612.

Christofides, N. and J.E. Beasley, "A Tree Search Algorithm for the P-Median Problem," *European Journal of Operations Research*, 10, 1982, 196-204.

Cohen, M.A., P.R. Kleindorfer and H.L. Lee, "Near-optimal service constrained stocking policies for spare parts," *Operations Research*, 37(1), 1989, 104-117.

Conn, A.R. and G. Cornuéjols, "A Projection Method for the Uncapacitated Facility Location Problem," *Mathematical Programming (A)* 46(3), 1990, 273-298.

Cooper, L., "Location-Allocation Problems," *Operations Research*, 11, 1963, 331-343.

Cooper, L., "Heuristic Methods for Location-Allocation Problems," *SIAM Review*, 6, 1964, 37-53.

Cornuéjols, G., "Analysis of Algorithms for a Class of Location Problems," working paper 382, School of Operations Research and Industrial Engineering, Cornell University, 1978.

Cornuéjols, G., M.L. Fisher and G.L. Nemhauser, "Location of Bank Accounts to Optimize Float: An Analytic Study of Exact and Approximate Algorithms," *Management Science* 23, 1977, 789-810.

Cornuéjols, G., G.L. Nemhauser and L.A. Wolsey, "A Canonical Representation of Simple Plant Location Problems and its applications," *SIAM Journal of Algebraic and Discrete Methods*, 1 (3), 1980, 261-272.-

Cornuéjols, G., G.L. Nemhauser and L.A. Wolsey, "The uncapacitated facility location problem," *Discrete Location Problems*, R.M Francis, and P.B. Mirchandani (eds), John Wiley and Sons, New-York, 1984.

Cornuéjols, G., R. Sridharan and J.-M. Thizy, "A Comparison of Heuristics and Relaxations for the Capacitated Plant Location Problem," *European Journal of Operational Research*, 50(3), 1991, 280-297.

Cornuéjols, G. and J.-M. Thizy, "Location problems, Set Covering Problems and the Greedy Algorithm," working paper 25-80-81, 1981, Graduate School of Industrial Administration, Carnegie-Mellon University, presented at the CORS-TIMS-ORSA Joint National Meeting, May 4, 1981, Toronto, Canada.

Cornuéjols, G. and J.-M. Thizy, "Some facets of the Simple Plant Location Polytope," *Mathematical Programming*, 23, 1982, 50-74.

Cornuéjols, G. and J.-M. Thizy, "A Primal Approach to the Simple Plant Location Problem," *SIAM Journal of Algebraic and Discrete Methods*, 3(4), 1982, 504-510.

Crainic, T.G., P. Dejax and L. Delorme, "Models for multimode multicommodity location problems with interdepot balancing requirements," *Annals of Operations Research*, Switzerland, 18(1), 1989, 279-302.

Current, J., H. Min and D. Schilling, "Multiobjective analysis of facility location decisions," *European Journal of Operational Research*, 49(3), 1990, 295-307.

Daskin, M.S., *Network and Discrete Location*, Wiley-Interscience, New York, 1995.

Davis, P.S. and T.L. Ray, "A Branch-and-Bound Algorithm for the Capacitated Facilities Location Problem," *Naval Research Logistics Quarterly*, 16(3), 1969, 331-344.

Dearing, P.M., P.L. Hammer and B. Simeone, "Boolean and Graph Theoretic Formulations of the Simple Plant Location Problem," *Transportation Science*, 26(2), 1992, 138-148.

Diehr, G., "An Algorithm for the p-Median Problem," Working Paper, No. 191, Western Management Science Institute, University of California at Los Angeles, June 1972.

Drezner, Z., "Location strategies for satellites' orbits," *Naval Research Logistics Quarterly*, 35(5), 1988, 503-512.

Drezner, Z., "Location of an obnoxious route," *Journal of the Operational Research Society*, 40 (11), 1989, 1011-1018.

Drysdale, J.K. and P.J. Sandiford, "Heuristic Warehouse Location - A Case History Using a New Method," *CORS Journal*, 7, 1969, 45-61.

Efroymsen, M.A. and T.L. Ray, "A Branch-and-Bound Algorithm for Plant Location," *Operations Research*, 14, 1966, 361-368.

Eilon, S. and R.D. Galvao, "Single and Double Vertex Substitution in Heuristic Procedures for the p-Median Problem," *Management Science*, 24(16), 1978, 1763-1766.

Eilon, S., C.D.T. Watson-Gandy and N. Christofides, *Distribution Management. Mathematical Modelling and Practical Analysis*, Griffin, London, 1971.

Ellwein, L.B. and P. Gray, "Solving Fixed Charge Location-Allocation Problems with capacity and configuration constraints," *AIIE Transactions*, 3(4), 1971, 290-298.

Elshafei, A.N., K.B. Haley and S.-J. Lee, "Facilities Location: Some Formulations, Methods of Solution, Applications and Computational Experience," The University of Birmingham, England, OR Report No. 90, 1974.

El-Shaieb, A.M., "A New Algorithm for Locating Sources Among Destinations," *Management Science*, 20, 1973, 221-231.

Erlenkotter, D., "Bibliography on Dynamic Location Problems," Discussion Paper No. 55, Management Science Study Center, Graduate School of Management, University of California at Los Angeles, 1975.

Erlenkotter, D., "A Dual-Based Procedure for Uncapacitated Facility Location," *Operations Research*, 26(6), 1978, 992-1009.

Erlenkotter, D., "An early classic misplaced: Ford W. Harris's economic order quantity model of 1915," *Management Science*, 35(7), 1989, 898-900.

Erlenkotter, D., "Ford Whitman Harris and the economic order quantity model," *Operations Research*, 38(6), 1990, 937-946.

Even, S., *Graph Algorithms*, Computer Science Press, Potomac, Maryland, 1979.

Feldman, E., F.A. Lehrer and T.L. Ray, "Warehouse Location Under Continuous Economies of Scale," *Management Science*, 12, 1966, 670-684.

Fisher, M.L. and D.S. Hochbaum, "Probabilistic Analysis of the Planar K-Median Problem," *Mathematics of Operations Research*, 5, 1980, 27-34.

Francis, R.M. and J.M. Goldstein, "Location Theory: A Selective Bibliography," *Operations Research*, 22, 1974, 400-409.

Francis, R.L., L.F. McGinnis, and J.A. White, "Locational Analysis," *European Journal of Operational Research*, 12, 1983, 220-252.

Francis, R.M. and P.B. Mirchandani, *Discrete Location Problems*, John Wiley and Sons, New-York, 1984.

Francis, R.M. and J.A. White, *Facility Layout and Location: an Analytical Approach*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1974.

Galvao, R.D., "A Dual-Bounded Algorithm for the p-Median Problem," *Operations Research*, 28(5), 1112-1121.

Gao, L.-L., "A dual-based optimization procedure for the two-echelon uncapacitated facility location problem," *Naval Research Logistics Quarterly*, 39(2), 1992, 191-212.

Garey M.R. and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, San Francisco 1979.

Garfinkel, R.S., A.W. Neebe and M.R. Rao, "An Algorithm for the M-Median Plant Location Problem," *Transportation Science*, 8, 1974, 217-236.

Geoffrion, A.M., "Lagrangian Relaxation for Integer Programming," *Mathematical Programming Study*, 2, 1974, 82-114.

Geoffrion, A.M., "A Guide to Computer-Assisted Methods for Distribution Systems Planning," *Sloan Management Review*, 16, 1975, 17-41.

Geoffrion, A.M., "Better Distribution Planning with Computer Models," *Harvard Business Review*, 54(4), 1976, 92-99.

Geoffrion, A.M. and G.W. Graves, "Multicommodity Distribution System Design by Benders Decomposition," *Management Science*, 20(5), 1974, 822-844.

Geoffrion, A.M. and R. McBride, "Lagrangian Relaxation Applied to Capacitated Facility Location Problems," *AIIE Transactions*, 10, 1978, 40-47.

Graves, S.C. "Determining the spares and staffing levels for a repair depot," *Journal of Manufacturing and Operations Management*, 1(2), 1988, 227-241.

Guignard, M., "Fractional vertices, Cuts and Facets of the Simple Plant Location Problem," *Mathematical Programming Study*, 12, 1980, 150-162.

Guignard, M. and K. Spielberg, "Algorithms for exploiting the Structure of the Simple Plant Location Problem," *Annals of Discrete Mathematics*, 1, 1977, 247-271.

Guignard, M., and K. Spielberg, "A Direct Dual Method for the Mixed Plant Location Problem with Some Side Constraints," *Mathematical Programming*, 17(2), 1979, 198-228.

Hadley, G. and T. Whitin, *Analysis of Inventory Systems*, Prentice-Hall, 1963.

Handler, G.Y. and P.B. Mirchandani, *Location on Networks: Theory of Algorithms*, M.I.T. Press, Cambridge, Mass., 1979.

Hansen P., "Two Algorithms for the Simple Plant Location Problem Using Additive Penalties," presented at the European Congress of the Econometric Society, Budapest, 1972.

Hansen P. and L. Kaufman, "An Algorithm for Central Facilities Location under an Investment Constraint," CS00TW/033, Centrum Voor Statistiek en Operationeel Onderzoek, Vrije Universiteit Brussel, 1972.

Harris, F., "Operations and Costs," *Factory Management Series*, Shaw Co., Chicago, 1915, 48-52.

Hochbaum, D.S., "Heuristics for the Fixed Cost Median Problem," *Mathematical Programming*, 22(2), 1982, 148-162.

Holmberg, K., "Experiments with Primal-dual Decomposition and Subgradient Methods for the Uncapacitated Facility Location Problem," Research Report LiTH-MAT-R-1995-08, Linköping Institute of Technology, Sweden, 1995.

Jack, N., "Costing a finite minimal repair replacement policy," *Journal of the Operational Research Society*, 43(3), 1992, 271-275.

Jacobsen, S.K. and P.M. Pruzan, "Lokalisering - modeller & losningsmetoder," IMSOR, Danmarks tekniske Højskole, 1976.

Jarvinen, P., J. Rajala and J. Sinervo, "A Branch-and-Bound Algorithm for Seeking the p-Median," *Operations Research*, 20, 1972, 173-178.

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

Karg, R.L. and G.L. Thompson, "A Heuristic Approach to Solving Traveling Salesman problems," *Management Science*, 10, 1964, 225-248.

Kaufman, L., "The Location of Economic Activities by 0-1 Programming," Vrije Universiteit Brussel, Brussel, 1975.

Khumawala, B.M., "An Efficient Branch-and-Bound Algorithm for the Warehouse Location Problem," *Management Science*, 18, 1972, B-718-731.

Khumawala, B.M., "An Efficient Heuristic Procedure for the Uncapacitated Warehouse Location Problem," *Naval Research Logistics Quarterly*, 20, 1973, 109-121.

Khumawala, B.M., "An Efficient Heuristic Procedure for the Capacitated Warehouse Location Problem," *Naval Research Logistics Quarterly*, 21(4), 1974, 609-623.

Khumawala, B.M., "An Efficient Branch-and-Bound Algorithm for the Capacitated Warehouse Location Problem," *Management Science*, 23(6), 1977, 585-594.

Khumawala, B.M., A.W. Neebe and D.G. Dannenbring, "A Note on El-Shaieb's New Algorithm for Locating Sources Among Destinations," *Management Science*, 21(2), 1974, 230-233.

Kochman, G.A. and C.J. McCallum Jr., "Facility Location Models for Planning Transatlantic Communications Network," *European Journal of Operational Research*, 6, 1981, 205-211.

Kolesar, P., "Testing for Vision Loss in Glaucoma Suspects," *Management Science* 26(5), 1980, 439-450.

Körkel, M., "On the exact solution of large-scale simple plane location problems," *European Journal of Operational Research*, 39(2), 1989, 157-173.

Krarp, J. and O. Bilde, "Plant Location, Set Covering, and Economic Lot-Size: An $O(mn)$ Algorithm for Structured Problems," in *Numerische Methoden bei Optimierungsaufgaben, Band 3: Optimierung bei Graphentheoretischen und Ganzzahligen Problemen*, Birkhauser, 1977, 155-186.

Krarup, J. and P.M. Pruzan, "Selected families of Discrete Location Problems. Part III: The Plant Location Family," Faculty of Business, University of Calgary, Calgary, Alberta, WP-12-77, Working Paper, 1977.

Krarup, J. and P.M. Pruzan, "Selected families of Location Problems," *Annals of Discrete Mathematics*, 5, 1979, 327-387.

Kuehn, A.A. and M.J. Hamburger, "A Heuristic Program for Locating Warehouses," *Management Science*, 9(4), 1963, 643-666.

Labbé, M. "The Median Location Problem with continuous demand," *Cahiers du Centre d'Etudes de Recherche Opérationnelle*, 34 (2/3/4), 1992, 103-115.

Lea, A.C., "Location-allocation Systems : an Annotated Bibliography," Discussion paper 13, Department of Geography, University of Toronto, 1973.

Lee, H.L. and K. Moinzadeh, "A repairable item inventory system with diagnostic and repair service," *European Journal of Operational Research*, 40(2), 1989, 210-221.

Levy, F.K., "An Application of Heuristic Problem Solving to Accounts Receivable Management," *Management Science*, 12, 1966, B-236-244.

Louveaux, F.V., "Stochastic location analysis," *Location Science*, 1 (2), 1993, 127-154.

Love, R.F., Morris, J.G. and G.O. Wesolowsky, *Facilities Location; Models and Methods*, North-Holland, Amsterdam, Netherlands, 1988.

Magnanti, T.L., "Combinatorial Optimization and Vehicle Fleet Planning: Perspectives and Prospects," *Networks* 11(2), 1981, 179-213.

Magnanti, T.L. and R.T. Wong, "Accelerating Benders Decomposition: Algorithmic Enhancement and Model Selection Criteria," *Operations Research*, 29, 1981, 464-484.

Manne, A.S., "Plant Location under Economies-of-Scale - Decentralization and Computation," *Management Science*, 11(2), 1964, 213-235.

Maranzana, F.E., "On the Location of Supply Points to Minimize Transport Costs," *Operational Research Quarterly*, 15, 1964, 261-270.

Markis, A. "A Generalized Multi-commodity Location-Allocation Model," *EIK*, 25 (7), 1989, 343-350.

Marsten, R.E., "An Algorithm for Finding Almost All the Medians of a Network," Discussion Paper, The Center for Mathematical Studies in Econometrics and Management Science, Northwestern University, No. 23, 1972.

Martel, S. and R.H. Bayne, "ASIM and a Naval Repair Problem," D Log A Staff Note 91/4, Department of National Defence, Ottawa, Canada, June 1991.

Mavrides, L.P., "An Indirect Method for the Generalized k-Median Problem Applied to Lock-Box Location," *Management Science*, 25(10), 1979, 990-996.

Megiddo, N., "Dynamic location problems," *Annals of Operations Research*, 6(4), 1986, 313-319.

Mirchandani, P.B. and R.L. Francis, eds., *Discrete Location Theory*, Wiley-Interscience, 1990.

Mirchandani, P.B., A.Oudjt and R.T. Wong, "Locational Decisions on Stochastic Multidimensional Networks," Rensselaer Polytechnic Institute, 1983.

Mirzaian, A., "Lagrangian Relaxation for the Star-star Concentrator Location Problem: Approximation Algorithm and Bounds," *Networks*, 15, 1985, 1-20.

Mulvey, J.M. and M.P. Beck, "A Heuristic-Subgradient Algorithm for Capacitated Clustering Problems," Princeton University Report EES-83-2, 1982.

Mulvey, J.M., and H.L. Crowder, "Cluster Analysis: An Application of Lagrangian Relaxation," *Management Science*, 25, 1979, 329-340.

Mulvey, J.M. and H. Vladimirov, "Stochastic network optimization models for investment planning," *Annals of Operations Research*, 20(6), 1989, 187-217.

Mulvey, J.M. and H. Vladimirov, "Applying the progressive hedging algorithm to stochastic generalized networks," *Annals of Operations Research*, 31(2), 1991, 399-424.

Narula, S.C., U.I. Ogbu and H.M. Samuelsson, "An Algorithm for the P-Median Problem," *Operations Research*, 25, 1977, 709-713.

Neebe, A.W. and M.R. Rao, "A subgradient approach to the m-median problem," University of North Carolina at Chapel Hill, 1975, 75-12.

Nemhauser, G., "The Age of Optimization : Solving Large-Scale real-world problems," *Operations Research*, 42(1), 1994, 5-13.

Ng K.Y.K., M.N. Lam and J.-M. Thizy, "Sensitivity analysis of the continuous review (Q,r) inventory-control system with service level constraint," Faculty of Administration, University of Ottawa, 1997.

Oral, M., M.S. Salvador, A. Reisman and B.V. Dean, "On the evaluation of shortage costs for inventory control of finished goods," *Management Science*, 18, 1972, B344-B351.

Pochet, Y. and L.A. Wolsey, "Lot-size Models with Back-logging: Strong Reformulations and Cutting Planes," *Mathematical Programming* 40, 1988, 317-335.

Pugliese, D.F. "Forces overstocked despite cuts," *The Ottawa Citizen*, Ottawa, 18 March 1996, 1.

Pyke, D.F. "Priority repair and dispatch policies for repairable-item logistics systems," *Naval Research Logistics Quarterly*, 37(1), 1990, 1-30.

ReVelle, C.S. and R.W. Swain, "Central Facilities Location," *Geographical Analysis*, 2, 1970, 30-42.

Rosing, R.E. and C.S. ReVelle, "A Method for Optimal Solutions to Large P-Median Problems," Economic Geographic Institute, Erasmus University, Rotterdam, 1978.

Rushton, Gerard, "Applications of location models," *Annals of Operations Research*, Switzerland, 18(1), 1989, 25-42.

Salkin, H.M., *Integer Programming*, Addison-Wesley, Reading, Mass., 1975.

Schilling D.A., V. Jayaraman and R. Barkhi, "A review of covering problems in facility location," *Location Science*, 1 (1), 1993, 25-55

Schneider, H., "Methods for determining the re-order point of an (s,S) ordering policy when a service level is specified," *Journal of the Operational Research Society*, 29, 1978, 1181-1193.

Schrage, L., "Implicit Representation of Variable Upper Bounds in Linear Programming," *Mathematical Programming Study*, 4, 1975, 118-132.

Schutt, J.H., "Mixed Integer Linear Programming Formulations for Distribution Networks," technical note, Cleveland Associates, Cleveland, 1979.

Simao, H.P. and J.-M. Thizy, "A Simplex Algorithm for the Canonical Representation of the Uncapacitated Plant Location Problem," *Operations Research Letters* 8(5), 1989, 279-286.

Simkin, L.P., "SLAM: Store Location Assessment Model-Theory and practice," *Omega*, 17(1), 1989, 53-58.

Spielberg, K., "Algorithms for the Simple Plant Location Problem with Some Side Conditions," *Operations Research*, 17, 1969, 85-111.

Spielberg, K., "Plant Location with Generalized Search Origin," *Management Science*, 16(3), 1969, 165-178.

Spielberg, K., "On Solving Plant Location Problems," *Applications of Mathematical Programming Techniques*, E.M.L. Beale ed., American Elsevier Publishing Inc., New York, 1970, 216-234.

Sridharan, R. "A Lagrangian Heuristic for the Capacitated Plant Location Problem with Single Source Constraints," *European Journal of Operational Research*, 66, 1993, 305-312.

Swain, R.W., "A Parametric Decomposition Approach for the Solution of Uncapacitated Location Problems," *Management Science*, 21, 1974, 189-198.

Teitz, M.B. and P. Bart, "Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph," *Operations Research*, 16, 1968, 955-961.

Thizy, J.-M., "Analysis of a new bound for the Greedy Heuristic applied to uncapacitated facility location," Faculty of Administration, University of Ottawa, 1997.

Thizy, J.-M., L.N. Van Wassenhove and B.M. Khumawala, "Comparison of Exact and Approximate Methods of Solving the Uncapacitated Plant Location Problem," *Journal of Operations Management*, 6(1), 1985, 23-34.

Van Hoesel, C.P.M., "Polyhedral results for the uncapacitated facility location problem: Effectiveness of inequalities and new facets," *Mathematical Programming Symposium*, Ann Arbor, 1994

Van Roy, T.J., "Cross Decomposition for Large-Scale, Mixed Integer Linear Programming with Applications to Facility Location on Distribution Networks," *Afdeling Industrieel Beleid*, Katholieke Universiteit Leuven, 1980.

Van Roy, T.J., "Cross Decomposition for Integer Programming," *Mathematical Programming*, 25(1), 1983, 46-63.

Van Roy, T.J., "A Cross Decomposition Algorithm for Capacitated Facility Location," *Operations Research*, 34(1), 1986, 145-163.

Van Roy, T.J. and D. Erlenkotter, "A Dual-Based Procedure For Dynamic Facility Location," *Management Science*, 28(10), 1982, 1091-1105.

Van Roy, T.J. and L.F. Gelders, "Solving a Distribution Problem with Side Constraints," *European Journal of Operational Research*, 6, 1981, 61-66.

Wong, R.T., "Accelerating Benders Decomposition for Network Design," PhD. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1978.

Wymore, A.W., *Model-based systems engineering*, CRC Press, Boca Raton, 1993.

Yano, C.A., "New algorithms for (Q,r) systems with complete backordering using a fill-rate criterion," *Naval Research Logistics Quarterly*, 32, 1985, 675-688.

Yano, C.A., "On the Equivalence of an equipment replacement problem and a facility location problem," Technical Report, Dept of Operations Engineering, University of Michigan, 1985.

Yip, H. "Level of Repair Analysis of an LRU from a high-frequency radio system," D Log A Project Report PR-9525, Directorate of Logistics Analysis, Department of National Defence, 1995.

Zheng, Y.S., "On properties of stochastic inventory systems," *Management Science*, 38, 1992, 87-103.