stichting
mathematisch
centrum

AFDELING MATHEMATISCHE BESLISKUNDE BW 117/80 FEBRUARI
(DEPARTMENT OF OPERATIONS RESEARCH)

A. FEDERGRUEN & B.J. LAGEWEG

HIERARCHICAL DISTRIBUTION MODELLING WITH ROUTING COSTS

Preprint

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O).

1980 Mathematics subject classification: 90B35

Hierarchical distribution modelling with routing costs **)

by

A. Federgruen*) & B.J. Lageweg

ABSTRACT

This paper describes modelling devices for hierarchical distribution systems, performing the distribution of commodities from producer to customer through a number of levels. For such a system we investigate *location* issues (number, location and capacity of facilities on each level), *allocation* issues (the allocation of delivery points to supply points on the next higher level), and in particular, *channel choices* (the level from which each delivery point should be served). Moreover the model assumes that several deliveries can be combined into a single route. This feature prohibits a simple allocation of distribution costs to individual demand points and necessitates explicit estimates of the *routing* costs into the model to be incorporated. Finally a description is given of an industrial project from which this paper originates and some computational experience with respect to this project is quoted.

KEYWORDS & PHRASES: *distribution system, hierarchical system, vehicle routing, location, allocation, multiple delivery journeys*

*) *Graduate School of Business, Columbia University, New York, U.S.A.*

**) This report has been submitted for publication elsewhere.

1. INTRODUCTION

1.1 A general hierarchical distribution system model

In recent years substantial efforts have been made to develop management science models enabling a comprehensive analysis of physical distribution systems. The hierarchy of a typical distribution system tends to consist of several layers; a number of layers between two and five is usual. The top level consists of the sources of the company's commodities: the production sites and/or the locations of the external suppliers. From here the dispatching of the company's commodities finds its origin, the final destination being the customer at the bottom level(s) of the hierarchy. Frequently these distribution processes use intermediate levels where the commodities can be stored, packed or repacked, and from where one may switch to a different and more economical mode of transportation. These intermediate facilities usually referred to as *distribution centers* or *depots* can therefore serve the following purposes:

- they provide regional buffer stocks;
- they station a fleet of company carriers adapted to the regional distribution characteristics;
- they enable assembling, repacking and bulkbreaking activities;
- they perform some of the later stages in the production process, e.g. when a trade-off of the production and distribution costs exhibits substantial advantages of shipping less voluminous or easier transportable halfproducts to these distribution centers;
- they represent marketing entities with specific regional expertise and canvassing and servicing responsibilities.

Another interesting feature shared by numerous distribution systems is the fact that the deliveries to some customers pass over one or more of the intermediate levels (cf. Figure 1, where part of the customers occur at the one but lowest level, obtaining *direct* deliveries from the second level in the network).

Frequently the distribution between at least some of the levels in the system is performed using *routes* with more than one delivery point. The incidence of trips involving multiple deliveries represents a major complication when evaluating the transportation costs for a particular distribution network. As long as delivery points are served on an individual basis the transportation costs can be represented by a separable function in the transportation flows. The routing aspect however adds a combinatorial element to the evaluation of the transportation costs. The marginal costs for a specific delivery point cannot be estimated from the mere knowledge of its location and order size; they can be insignificant if a vehicle visits the neighbourhood of the point with sufficient spare capacity and time; on the other hand an extra route or even an extra vehicle may be involved if no spare capacity is available on anyone of the existing routes. Reports [MINISTRY OF TRANSPORT 1959, 1964, 1965] point out that the vast majority of all road good vehicles in Britain are used on multiple delivery routes and 80% of the latter are performed by company owned vehicles. Similar data have been reported for other countries [BAYLISS 1965; UNITED NATIONS 1967; SAMPSON & FARIS 1966].

Planning studies for these distribution systems may involve a variety of questions:

- *network depth*: the number of levels in the system;
- *location issues*: the number and locations of facilities on each level;
- *allocation issues*: the allocation of delivery points to supply points on the next higher level;
- *product mix*: the products each of the facilities should carry;
- *capacity issues*: the production and throughput capacities in each facility;
- *modes of transportation*: the modes of delivery to be used for the various transportation flows;
- *channel choices*: the level from which each customer should be served.

A comprehensive planning model is needed when analyzing the design of a new system, or when undertaking a major reorganisation of an existing one. More frequently one faces specific questions with respect to a small number of facilities or customers. These questions, although of a seemingly regional or individual character, have to be addressed in the context of the national distribution network. We refer to Geoffrion [GEOFFRION 1976] for an elegant exposition of the use of computer guided models for distribution system planning.

This paper shows how a number of the above listed problems can be incorporated into a management science model with an explicit and adequate representation of the transportation costs on "multiple delivery routes". We use a specific model to explain our approach. Both the characteristics of the distribution system and the specific set of problems in this model were motivated by a study undertaken for AGA BV, a producer and distributor of industrial gases in the Netherlands. This application is discussed at some length in section 6, where computational experience will be given.

We conclude this introduction by a quotation [HERRON 1979] on the importance of an efficient Physical Distribution Management (PDM): "Many companies tend to shrug off PDM on the assumption that distribution effectiveness does not make much difference anyway. Yet recent studies have shown that both the cost and income effects of distribution are surprisingly large... (and) that the physical distribution of goods from producer to final consumer costs more than $400 billion a year, or 20% of GNP." The same source mentions that distribution costs average 13.6% of the sales dollar for manufacturing companies and 25% for merchandising companies, with direct labour comprising a mere 12-14% share, and concludes that "PDM evidently represents a considerably greater potential for productivity improvement than does the much discussed labour output per man-hour".

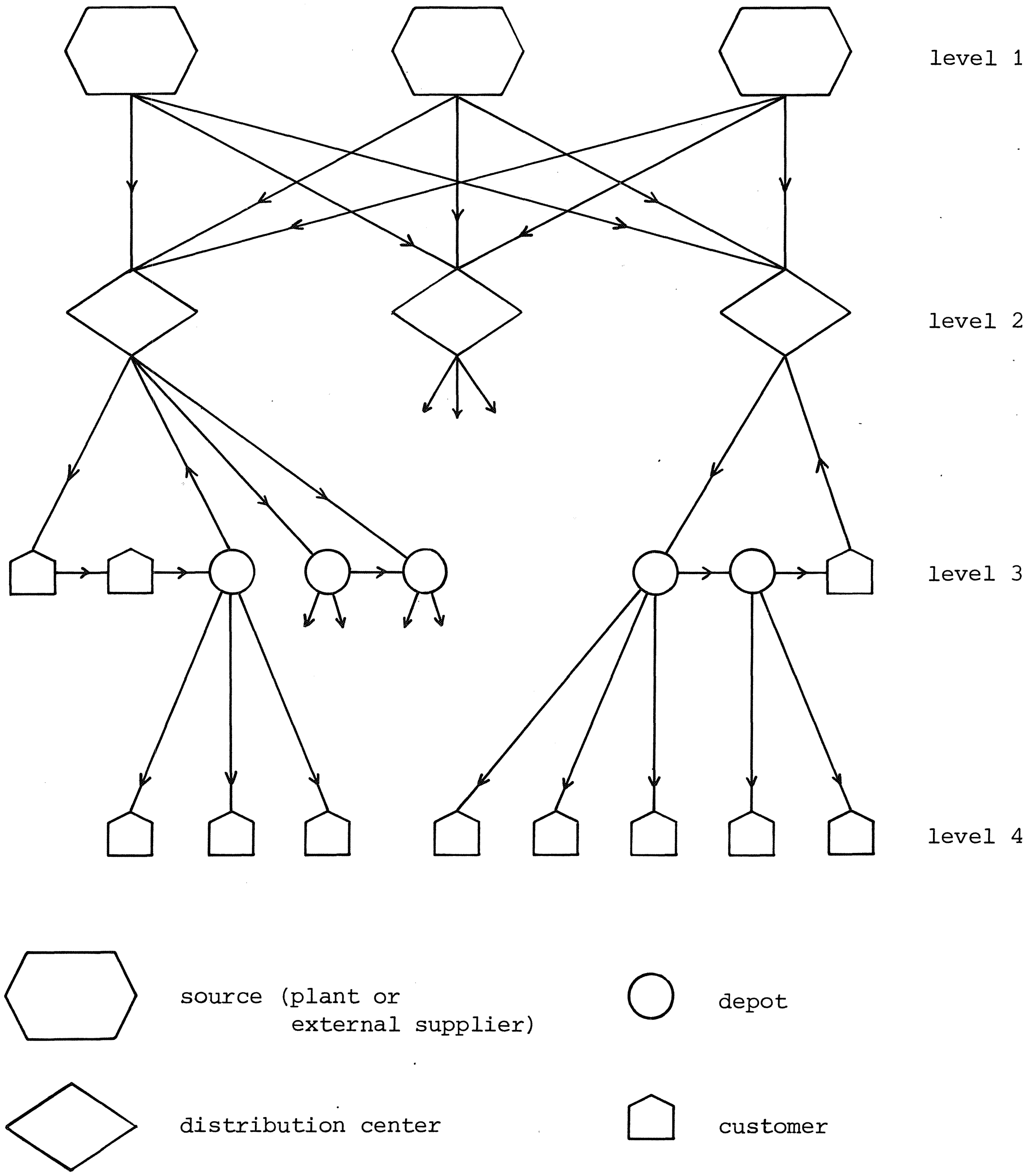
level 1
level 2
level 3
level 4
source (plant or
external supplier)
depot
distribution center
customer

Figure 1

1.2 Outline of the paper

In section 2 we define a special hierarchical distribution system as well as the network design problems that we wish to approach. We next discuss how this model relates to the existing literature and explain why a different modelling approach seems unavoidable in our case. Section 3 discusses the variables of the model which specify the design of the network, and their quantification. The evaluation of the various cost components for a given design of the network is described in Section 4. We especially emphasise the problem of estimating the transportation costs between intermediate levels in the network where multiple deliveries are combined into routes and where the destinations have different frequencies of delivery. We discuss procedures to characterize the typical workload in a standard period of (say) a week, and present heuristic solution methods for the resulting routing type problems. Section 5 is devoted to optimization procedures for the variables that specify the design of the network. The above mentioned industrial project from which this model has arisen is described in Section 6. Finally, three appendices deal with some technical and numerical questions, such as the computation of customer characteristics and the calculation of distances between customer and facility locations.

2. A SPECIFIC MODEL

A company produces and distributes several commodities. The distribution process originates at a number of *sources* (cf. the top level in Figure 1) representing either company owned plants or external suppliers. We assume that for each source lower and upper limits are given for each commodity on the amount to be produced or supplied.

The dispatching of the commodities to their final destinations, the customers, is initiated via a number of national *distribution centers* (DC). These DC's perform some or all of the activities listed in the introduction and are represented at the second level of Figure 1. Especially with respect to the fourth listed function, the transformation of half-products into final products, we specify for each DC and each commodity a set of sources from which the commodity is only to be dispatched as a halfproduct, the complementary set of sources acting as potential suppliers of the commodity in its final form. In each DC a fleet of identical vehicles is stationed in order to forward the commodities to some of the customers as well as to *regional depots* which supply the remaining clients. The customers at the third level of the network with direct deliveries from one of the DC's will be denoted as *direct customers*, as opposed to the *indirect customers* receiving service via one of the regional depots. We conclude that in our specific model a network of four layers is considered with the final destinations occurring at the third and fourth level.

Individual dropsizes are such that the vehicles at the DC's usually combine several deliveries into a single route. In order to facilitate the presentation of our model, we assume that these *multiple delivery routes* only occur between the second and third level of the system. The individual service assumption is satisfied most times with respect to the flows of bulk transportation between the first and second level; moreover our approach can easily be adapted for the case where the final stage of the distribution process between regional depots and indirect customers is characterised by multiple delivery journeys as well, albeit at a considerable increase of the computational effort involved. Due to the use of different types of equipment e.g., transportation costs per time unit between regional depots and indirect customers are lower than the corresponding costs for a DC-stationed truck. This provides one rationale for the insertion of regional depots on trajectories to smaller or inconveniently located customers.

Customers receive their goods on a periodic basis. The frequency of delivery however can range from once a day up to say once a month. Both average dropsize and frequency of delivery are determined by the customer and are assumed to be given and fixed. Potential control of these quantities by transportation fees, drop charges, package rental fees etc. will not be considered in this model.

Each customer and regional depot is assumed to have a single source of supply for all its commodities. In case some diversity of supply sources is allowed, extensions of our approach can be achieved along the lines described by Geoffrion et al. [GEOFFRION et al. 1978]. Moreover, at the DC's all commodities are repacked into standard packing units. As a consequence the distinction between commodities is only needed for the flows of bulk transportation at the first stage of the distribution process.

2.1 The cost structure

The following cost components are considered.

1. *production/purchasing costs*: these occur at the company owned plants and at the location of the external suppliers and can be represented as functions of the volumes in the commodities concerned.

2. *throughput/production costs in the DC's*: this component is a function of the number of standard packing units to be dispatched via this particular DC. If the DC undertakes the final stage(s) of the production process of some commodities, the production volumes of those commodities occur as additional arguments in this function. We merely assume that this function is jointly concave in all arguments reflecting economies of scale. In the case study to be described in section 6, these functions are nonseparable in the arguments and concavity is brought about both by the occurrence of fixed costs and by nonlinearities in the variable costs. Note that this cost component may include inventory carrying costs for the commodities themselves and for the packing units used. We assume average inventories to depend on the annual throughput according to given concave functions.

3. *throughput costs in the regional depots*: for each regional depot these costs can be represented as a function of total throughput measured in standard packing units. Here again inventory carrying costs may be included.

4. *costs of bulk transportation*: due to the assumption of individual service of the DC's these costs are fully specified by a single unit inbound transportation rate for each commodity and each inbound

link, directed from a source into a DC.

5. *costs of direct deliveries*: these concern the deliveries between the DC's and the third level of the system. For the sake of notational simplicity, we assume all vehicle costs to be identical and independent of the location of the DC, where they are stationed. This component is specified by an annual amount of fixed costs per vehicle, a variable amount per unit time spent on the road or loading/unloading units, and finally a variable amount per mile. Moreover vehicle capacity restrictions are represented by upper bounds on the number of packing units and the length or duration of a route.

6. *costs of indirect deliveries*: the costs of delivery from a regional depot to a customer are assumed to be proportional to the flow between this pair, as a consequence of the assumption of individual service for the final stage of the distribution process.

We note that the first three cost components frequently require a one-time though conscientious exercise in operational accounting. Once these components have been established as functions of the various distribution and production volumes, their evaluation is straightforward, as is the case for components 4 and 6. On the contrary the evaluation of component 5 in many cases constituting a major or even dominant share of the total costs requires the solution of complex routing type problems.

2.2 The basic questions

The purpose of this paper is to develop an approach for the simultaneous analysis of the following interdependent questions.

1. Which of the list of potential distribution centers should be used.

2. Which order of magnitude should the number of regional depots have.

3. Which territory should each DC have, i.e. which regional depots and

which customers should be allocated, either directly or indirectly, to this DC.

4. Which customers should be given direct service from one of the DC's and which customers should obtain their deliveries from one of the regional depots. In other words, which customers should be put at the third and which at the bottom level of the distribution network (cf. Figure 1).

5. How should the annual transportation flows be organized throughout the system.

With respect to the problems under 4, the purpose is to develop a rule which allocates the customers to the direct or indirect channel on the basis of a limited number of customer characteristics. This provides an alternative to determining the allocations via a huge set of binary variables in some mixed integer programming formulation. Determining a *channel rule* (CR) as a function of customer characteristics thus accomplishes an immense reduction of the computational complexity; besides it has the advantage of creating insight into the rationales behind the allocations of the customers and of being applicable for future use either with respect to new customers joining the system or with respect to customers with a changed demand patter. In the case study to be described below, it was this problem area that ranked highest among management's priorities, and that a posteriori accounted for most recommended savings.

In addition we stress the difference in the formulation of the first two problems both of which concern location-type issues of transit points. As the number of regional depots usually is at least one order of magnitude larger than the number of DC's, one contents oneself frequently to the problem of selecting the desired density of the network of regional depots, instead of a precise set of locations as is required for the more capital and labor intensive DC's.

In the remainder of this paper we consider the problem of determining a design of the distribution network that minimizes the total costs, the components of which were enumerated in section 2.1. We ignore for the moment potential impacts on the revenue side, e.g. via the perceived service level. As another case in point, note that a decision to reduce the number of regional depots, although possibly cost efficient, may entail important but hardly quantifiable marketing consequences, such as a loss of customers with traditional and personal ties to the depot holders of the depots that are to be closed. As Geoffrion [GEOFFRION 1976] we emphasize that our model should be used as a mere decision support tool which has to be integrated with nonquantifiable elements like the above mentioned considerations.

2.3 Relation to literature

A few remarks should be made about the relation of our model to the existing literature.

The first three problem areas have the character of location/allocation problems, a subject with a rich background in OR-literature. We refer to bibliographies by Golden et al. [GOLDEN et al. 1977] and Francis and Goldstein [FRANCIS & GOLDSTEIN 1974] and to surveys by Ellwein [ELLWEIN 1970] and Eilon et al. [EILON et al. 1971]. Most of the literature has concentrated upon the single commodity transportation problems with fixed charges for the use of a source, either with or without capacity constraints on the throughput of sources. Extensions of this model to multilevel hierarchical systems were initiated by Marks et al. [MARKS et al. 1970]. Geoffrion et al. [GEOFFRION 1971; GEOFFRION 1976; GEOFFRION et al. 1978] recently developed a much valued and fairly general model for multicommodity hierarchical distribution systems, as we consider in this paper. The model is based on the classical assumption that delivery points receive individual service rather than in multiple delivery routes, or that customers somehow have been grouped into zones, with each zone receiving individual service. This assumption allows for a mixed integer programming formulation which can be solved by Benders' decomposition, albeit at the price of repeatedly solving an integer master program with several hundreds or thousands of integer variables.

We explained how the occurrence of multiple delivery routes complicates the representation of transportation costs and hence the analysis of changes in the design of the network. In view of the abundance of multiple delivery routing, as mentioned in section 1, surprisingly few attempts have been made to integrate location and routing models [WATSON-GANDY & DOHM 1973; TFD 1977; JACOBSEN & MADSEN 1978]. Christofides and Eilon [CHRISTOFIDES & EILON 1969] and Webb [WEBB 1968] were the first to advocate an adequate representation of routing aspects in location/allocation models. Webb has shown that simplified cost functions, e.g. linear or separable functions in the transportation flows, may lead to poor measures for the variable costs and to misleading results in the location area.

As far as vehicle routing is concerned, the last years have seen a substantial development of algorithms for one-period vehicle routing problems [EILON et al. 1971; CHRISTOFIDES 1976; GOLDEN et al. 1977; FISHER & JAIKUMAR 1978]. Only recently however a start has been made with more complex cases, where customers have different frequencies of delivery, and are to be assigned not only to routes but also to days of (say) a week, in such a way as to enable effective routing of the resulting one-day problems [BELTRAMI & BODIN 1974; RUSSEL & IGO 1979].

3. SPECIFYING THE DESIGN OF THE DISTRIBUTION NETWORK

In this section we discuss the variables that specify the structure and design of the distribution network, in particular as far as the five basic questions of section 2.2 are concerned:

1. the set of location sites for the DC's;
2. the set of location sites for the regional depots;
3. the allocation of the direct customers and the regional depots to one of the DC's, and the allocation of indirect customers to one of the regional depots, i.e. the specification of the territories of the DC's and the regional depots;
4. the *delivery mode* or *channel* for the customers, i.e. the allocation of the customers to the direct or indirect channel;
5. the bulk transportation flows for each commodity between the various sources-DC pairs.

First we introduce some notation.

3.1. Notation

Subscripts

h	indexes commodities	$(1 \le h \le H)$
i	indexes sources	$(1 \le i \le I)$
j	indexes DC's	$(1 \le j \le J)$
k	indexes regional depots	$(1 \le k \le K)$
ℓ	indexes customers	$(1 \le \ell \le L)$

Flow variables

x_{hij} : the amount of commodity h flowing from source i to DC_j, in standard packing units per year (spu/yr).

x_{hi}^{out} : the amount of commodity h produced or purchased at source i (spu/yr).

$x_i^{out} = (x_{1i}^{out}, \ldots, x_{hi}^{out}, \ldots, x_{Hi}^{out})$

x_{hj}^{in} : the amount of commodity h dispatched to DC_j (spu/yr).

ζ_{hj} : the amount of commodity h, undergoing a second stage(s) of production in DC_j (spu/yr).

$\zeta_j = (\zeta_{1j}, \ldots, \zeta_{hj}, \ldots, \zeta_{Hj})$

y_j : the amount of commodities distributed via DC_j (spu/yr).

z_k : the amount of commodities distributed via regional depot k (spu/yr).

Coefficients

c_{hij}^{bulk} : unit bulk transportation cost (\$/spu) for shipments of commodity h from source i to DC_j.

$c_{k\ell}^{dep}$: unit transportation cost (\$/spu) for shipments from regional depot k to customer ℓ.

SLO_{hi}, SUP_{hi}: lower, resp. upper limit on the total production/supply of commodity h at source i (spu/yr).

DLO_j, DUP_j : lower, resp. upper limit on the total throughput of DC_j, if it is open (spu/yr).

Q : capacity of a vehicle (spu).

T : maximum duration of a vehicle route for DC-stationed vehicles.

vf : fixed cost (\$/yr) of a DC-stationed vehicle.

vt : variable cost (\$ per unit time) of a DC-stationed vehicle.

vm : variable cost (\$ per mile) of a DC-stationed vehicle.

Cost functions

$\Theta_i(x_i^{out})$: the production or purchasing costs (\$/yr) at source i, given a supply vector x_i^{out}.

$\Phi_j(\zeta_j, y_j)$: the total production and throughput costs (\$/yr) at DC_j, given a production vector ζ_j and a distribution volume y_j.

$\Psi_k(z_k)$: the total throughput costs (\$/yr) at regional depot k, given a throughput volume z_k.

Balance equations

(3.1) $$x_{hi}^{out} = \sum_{j=1}^{J} x_{hij} \qquad (1 \le h \le H;\ 1 \le i \le I)$$

(3.2) $$x_{hj}^{in} = \sum_{i=1}^{I} x_{hij} \qquad (1 \le h \le H;\ 1 \le j \le J)$$

(3.3) $$y_j = \sum_{h=1}^{H} x_{hj}^{in} \qquad (1 \le j \le J)$$

(3.4) $$\zeta_{hj} = \sum_{i \in S(h,j)} x_{hij} \qquad (1 \le h \le H;\ 1 \le j \le J),$$

where S(h,j) is the set of sources which supply commodity h to DC_j as a halfproduct.

3.2 Design variables

The quantification of the sets of variables (1), (2) and (5), defined in the beginning of this section, is more or less straightforward.

With respect to (1) a list of potential locations for DC's needs to be specified, together with the cost parameters that depend upon these locations. We assume that any combination of DC's of this list of potential locations is allowed except for possible lower and upper limits on the number of DC's.

With respect to (2) we specify a list of regional depot sets, i.e. instead of allowing all subsets of a collection of potential depot locations we restrict our choice to a limited list of depot sets. This different specification for (2) is due to the fact that in case of the regional depots we are only interested in the order of density of the depot net, for reasons mentioned in Section 2.2.

For the bulk transportation flows (5) the flow variables x_{hij}, $1 \le h \le H$, $1 \le i \le I$, $1 \le j \le J$, will be used. The specification of the sets of allocation variables (3) and (4) needs somewhat more care.

We emphasized before that the channel choice problem should be modelled in such a way as to generate a *channel rule* which determines the allocation of customers to the direct or indirect channel on the basis of a limited number of customer characteristics. The remainder of this section will be devoted to the problem of selecting a set of relevant customer characteristics and their quantification, and to the selection of a class of parametrized channel rules. Note that by this construction an efficient channel rule can be found merely by searching for the "best" combination of paremeters or *critical values* within the aforementioned class.

This leaves us with the allocation variables in (3). The occurrence of multiple delivery routes, frequently cited malefactors, makes it difficult to represent the consequences of the allocations of individual direct customers and regional depots in a simple and tractable objective function. Accordingly we will present two distinct approaches to the problem of locating DC's and determining their territories:

- a model with allocation variables for each individual customer and depot to the various DC's;

- a model with allocation variables for clusters of customers; each cluster corresponds with a traveling salesman tour as generated by a vehicle routing algorithm.

The customers, that a channel rule determines to receive indirect deliveries, are assigned to the regional depot that within the depot set under consideration is closest in travel time or costs.

3.3 Customer characteristics

When considering the qualities of a customer that determine the desirability of providing him direct service, it is evident that the size of the customer's orders is a first candidate. To characterize the customer's order size we choose his *average dropsize* δ, i.e. the average number of standard packing units delivered. Note that no distinction between commodities is needed since all transportation costs incurred beyond the second level of our system are assumed to be identical for all commodities.

The customer's own demand is not the only relevant factor for a channel choice. In addition we would like to base the choice upon the demand pattern in the proximity of the customer's locations. Especially for customers with moderate dropsizes the desirability of providing direct service increases as more demand has to be satisfied within the customer's proximity. Whereas the relevancy of this characteristic is immediate, its quantification is far less obvious. Measurements via say the total demand within a predetermined radius around the customer's location arouse a multitude of questions and objections: how should the radius be chosen, does it matter whether the total demand is spread over a large number of customers or concentrated on a few locations: should the distribution of travel times to the customers within this radius be recorded as a separate characteristic?

To avoid this arbitrariness we developed an altogether different measure, the *degree of separation* σ of the customer, defined as the minimal transportation costs on a Hamilton path, which starts at the customer under consideration and next visits other customers until the number of spu's delivered on this route equals or exceeds the capacity of a truck. Here we fix the dropsizes of the various demand points at their expected or average values. A complicating feature is the variance in the customer's frequency of delivery: a customer with a high frequency should have a proportionally higher impact on the value of the measure compared to a customer with the same dropsize and a lower frequency. To achieve this we simulate the workload on a typical day by creating a series of so-called *standard days*.

A customer appears in a standard day with probability p_i equal to his average or expected number of deliveries per working day. In other words the workload during a standard day is determined by applying a random generator independently to each one of the customers with p_i representing customer i's probability of entering the standard day (see also Appendix 1).

With these conventions, the separation degree for each customer can be calculated as the average of the values of this measure over all standard days. Computation of the degree on a particular standard day requires the solution of a combinatorial optimization problem which can be considered as a generalization of the classical traveling salesman problem. Appendix 1 gives a mathematical programming formulation of this tour selection problem and a brief description of a number of heuristic and exact solution methods. In section 6 a map of the Netherlands will be shown, exhibiting the score of the various customer locations with respect to the separation degree.

We finally observe that this measure succeeds fairly well in capturing the routing costs in the proximity of the customer given the constraint that an a priori and one-time computable measure is wanted, into which no variable information regarding locations of DC's regional depots or allocation variables can be incorporated. Moreover, we emphasize that the measure merely serves the purpose of ranking the customers on a separation degree scale, and that the measure is not used to eliminate a cost component of the distribution system itself. This and the fact that the underlying tour selection problem has to be solved for every single customer justify the use of elementary heuristic solution methods for this combinatorial optimization problem.

Next to the average dropsize and the separation degree, other characteristics may be considered. In the aforementioned case study we experimented e.g. with a number of distance measures, such as the distance of the customer to the nearest DC and the distance to the nearest regional depot. However we observed that a customer who is located at a large distance from the nearest DC (say) is sometimes a better, and sometimes a worse candidate for direct service as compared to a customer which is closer to the DC and has otherwise identical characteristics.

The lack of these monotonicity properties which hold for the average dropsize and separation degree measures may exclude a simple structure of any channel rule based on such distance measures. As these measures moreover depend on variable information, i.e. have to be adapted whenever the set of DC's or regional depots is changed, we discarded them as customer characteristics. As a consequence we have only considered channel rules that are based on the above two measures δ and σ. An additional motivation for this choice was our desire to generate an easily implementable channel rule, and to reduce the number of degrees of freedom of the class of channel rules within which an optimal rule has to be found.

3.4 A class of channel rules

We consider a class of channel rules which determine the allocation of customers to the direct or indirect channel on the basis of the average dropsize δ and the separation degree σ. The class of rules has three parameters or threshold values δ_-, δ^+ and σ^+, and has the following structure:

- customers with $\delta < \delta_-$ are assigned to the indirect channel;
- customers with $\delta \geq \delta^+$ are assigned to the direct channel;
- for intermediate customers ($\delta_- \leq \delta < \delta^+$) the channel is determined by the separation degree: customers within this category for which $\sigma \leq \sigma^+$ will be provided with direct service, all others are assigned to the indirect channel.

Note that the structure of this rule allows for any possible combination of the relative impacts of the two characteristics, depending on the length of the interval (δ_-, δ^+). Rules with $\delta_- = \delta^+$ base the channel allocation entirely upon the average dropsize; conversely for a choice of $\delta_- = 0$ and $\delta^+ = \infty$ only the separation degree matters.

4. EVALUATING THE COSTS FOR A GIVEN DESIGN OF THE DISTRIBUTION NETWORK

In this section we describe how the various cost components enumerated in Section 2.1 can be evaluated for a given design of the distribution system, i.e. for a given choice of the five sets of variables specified in Section 3. We estimate the yearly costs from the costs incurred in a *standard period* which is defined as a sequence of (say) W consecutive days.

Let P denote the set of delivery points at the third level of the system, the regional depots and the direct customers. Furthermore we define

$A_{h\ell}$ = annual turnover of commodity h for customer ℓ (spu/yr)

$$A_\ell = \sum_{h=1}^{H} A_{h\ell}$$

δ_p = dropsize of demand point p (spu)

f_p = frequency of delivery of demand point p, i.e. the average number of deliveries per standard period, a rational number between 0 and W, where

W = length of the standard period.

Knowledge of the set of regional depots and the channel rule as specified by the triple $(\delta_-, \delta^+, \sigma^+)$ enables us to determine P, as well as (δ_p, f_p) for all $p \in P$. For a given depot the annual throughput can be computed by totalizing the annual demand A_ℓ of all indirect customers for which this regional depot is the closest among all depots in the set under consideration (cf. Section 3). A given inventory rule determines the frequency and hence the dropsize as a function of the depot's annual throughput.

We next verify that the evaluation of all cost components except the costs of direct deliveries is straightforward when knowing the values of the five sets of design variables:

1. production/purchasing costs: the costs in source i are given by $\Theta_i(x_i^{out})$, where x_i^{out} can be determined from the bulk flow transportation variables $\{x_{hij}\}$ via (3.1);

2. throughput/production costs in the DC's: the costs in DC_j are given by $\Phi_j(\zeta_j, y_j)$, where the production vector ζ_j and the distribution volume y_j can be determined from $\{x_{hij}\}$ via (3.2) - (3.4);

3. throughput costs in the regional depots: for depot k these are given by $\Psi_k(\Sigma_{\ell \in P_k} A_\ell)$, where P_k denotes the set of indirect customers assigned to depot k;

4. costs of bulk transportation: $\sum_h \sum_i \sum_j c_{hij}^{bulk} x_{hij}$;

6. costs of indirect deliveries: $\sum_k \sum_{\ell \in P_k} c_{k\ell}^{dep} A_\ell$.

This leaves us with the determination of cost component 5, the costs of direct deliveries, which is substantially more complicated due to the multiple delivery routing aspect. In view of the variability in the frequency of delivery among various customers and depots we first need to characterize the workload in a typical standard period. Next we face the problem of estimating the routing costs for a given workload.

4.1. The workload in a standard period

We consider a period of W consecutive working days, e.g. in the case study below a standard week from Monday to Friday with W = 5. The frequency $\hat{f}_p$ with which a demand point $p \in P$ appears in this standard week, is determined by rounding off the usually non-integer value f_p to $\lceil f_p \rceil$ with a probability given by its residual part and to $\lfloor f_p \rfloor$ with the complementary probability. For instance, a customer p with $f_p = 2.7$ has a probability of 70% of receiving three deliveries in the standard week and a probability of 30% of only two deliveries. Therefore not all direct customers will appear in the standard week: a direct customer p with $f_p < 1$ has probability $1-f_p$ to be left out.

Note that the creation of these standard periods is based on a similar procedure as the generation of standard days for the computation of the separation degree (cf. Section 3.3). Here the underlying assumption is that the frequency of a customer's deliveries is a random variable with a distribution concentrated on the two integers closest to its expected value and independent of the frequency of delivery of other customers. This assumption obviously underestimates to some extent the variance and the pairwise correlation coefficient of these random variables (cf. Appendix 1 for a discussion of the latter phenomenon).

We finally note that for all customers the standard frequency $\hat{f}_\ell$ and the dropsize δ_ℓ can be determined without specifying any of the design variables of the distribution system. However to determine the set of demand points $\hat{P}$ in a standard period and their standard frequency $\hat{f}_p$ and dropsize δ_p we need to specify the channel rule and the set of regional depots. A customer ℓ which is to receive direct service by a given channel rule will appear in the standard period if $\hat{f}_\ell \geq 1$. The remaining, indirect customers contribute all to the annual throughput of the regional depot to which they are allocated and thus determine its standard frequency and dropsize.

4.2 Costs of direct deliveries in a standard period

This subsection deals with the problem of estimating the vehicle routing costs associated with the direct deliveries between DC's and the set $\hat{P}$ of direct destinations in the standard period. Knowledge of the territories of the DC's (cf. the third set of variables in section 3) enables a decomposition of the problem into several *one-center* multi-period vehicle routing problems (MVRP). In this generalization of the classical one-center *one-period* vehicle routing problem (VRP) demand points are to be assigned to a given number of days in a standard period of W days so as to minimize the sum of the resulting W VRP-costs.

The simplest version of the MVRP specifies the following constraints:

- an upper limit Q on the vehicle capacity, assuming that only one type of vehicle is to be used;
- an upper limit T on the duration of a route, defined as the sum of the travel times on the links belonging to the route plus the sum of loading/unloading times at the delivery points. The former are computed by determining shortest paths in a road network (see Appendix 3), the latter are found from a loading time function which may depend on the (type of) location and the dropsize of the delivery point;
- a set of *spacing conditions* indicating which combinations of days are allowed for the various demand points. In principle and primarily for operational day to day scheduling problems these conditions may specify any collection of permitted day combinations for each demand point. However in the context of strategical studies only simple and general spacing conditions for an entire class of demand points with a specific frequency are likely to be considered, such as a minimum or maximum number of days between consecutive deliveries. E.g. for a demand point p with $\hat{f}_p = 2$ the spacing conditions may merely forbid deliveries on consecutive days.

In addition to the constraints above a multitude of extra complications may arise, as in the simple VRP-model: several types of vehicles, multiple commodity compartments, various capacity measures, time windows within which deliveries have to occur. Once again these complications are unlikely to occur in tactical studies.

As in the case of simple VRP's several objectives may be minimized [CHRISTOFIDES 1976]. If the vehicle costs are mainly fixed costs such as vehicle investment costs and (fixed) drivers' salaries, one may try to find a lexicographic minimum to the following two objectives.

- the number of vehicles, which is given by the maximum number of vehicles required on any day of the standard period;

- the variable routing costs which usually have a time and a mileage dependent component (cf. the cost parameters vt and vm in section 3.1).

As an alternative one may minimize the total costs function, i.e. the sum of the fixed and variable costs. Especially in the former case we observe that a good assignment of customers to the days of the standard period seems to satisfy the property of equalizing as much as possible over those days the number of vehicles in use and secondarily the total daily routing costs.

Algorithms to solve the MVRP are not abundant in the literature. Whereas considerable progress has been made in the study and implementation of exact and heuristic methods for the one-day VRP [EILON et al.1971; CHRISTOFIDES 1976; GOLDEN et.al 1977; FISHER & JAIKUMAR 1978] only the work of Beltrami and Bodin [BELTRAMI & BODIN 1974] and recently Russell and Igo [RUSSEL & IGO 1979] are to be mentioned in the context of MVRP's. The former develop a heuristic for the special case with W = 6 and $\hat{f}_p = 3$ or $\hat{f}_p = 6$. Russell and Igo have experimented with a number of heuristics most of which extend the ideas of classical algorithms [CLARKE & WRIGHT 1964; LIN & KERNIGHAN 1973] to the MVRP case.

Most approaches of the MVRP require the construction of an initial feasible assignment of the customers to the days of the standard period. With respect to the second stage of the algorithm two types of approaches have been suggested.

1. One estimates $R_i, i = 1,\dots,W$, the number of routes on each one of the W days. The initial assignment of customers to days is merely used to select for each day and for each route to be driven on that day a so-called *seed customer*. Finally the remaining customers are assigned to these routes by solving some generalized assignment problem or by applying a savings heuristic. This procedure is repeated for a number of estimates of R_i, $i = 1,\dots,W$;

2. The initial customer-to-days assignment is used to solve each one on the resulting W VRP's. Next, one may consider feasible interchanges of customers among routes belonging to different days, thus

creating different customer-to-days assignments. This approach in fact was taken in the case study described in Section 6, where the VRP-package VERSA [CHRISTOFIDES 1975] was used to solve the VRP's.

In both approaches the success of the method depends heavily on the availability of a good heuristic for the initial customers-to-days assignment. We consider the study of this problem as one of the immediate challenges in the vehicle routing field. Appendix 2 gives a description of the heuristic utilized in our case study.

5. THE OPTIMIZATION OF THE DESIGN VARIABLES

In this section we present several procedures for the optimization of the five sets of design variables given in Section 3. In view of the overwhelming complexity of the entire problem it is obvious that some hierarchy among the sets of variables has to be established, so that

- a lower level optimization model finds the optimal values of the "lower" variable sets for a given choice of values of the "higher" variable sets;
- this lower level model is used as a subroutine in a procedure which searches for an optimal combination of higher level variables.

Obviously a repeated use of this *nesting* device may be needed, i.e. one may wish to establish a similar hierarchy within the lower level variables, etc. We suggest the following nesting device:

1. designate the channel rule and the set of regional depots as the higher level variables (sets 2 and 4);
2. develop a location-allocation-routing model (ALLOCRO) which determines an optimal combination of the remaining lower level variable sets (sets 1,3 and 5), i.e. the sets of DC's, their territories and the bulk flow variables, for any given choice of channel rule and regional depot set, and estimate all cost components (cf. Section 4) for this design.

In Section 3 we observed that with respect to the regional depots a selection has to be made within a predetermined list of collections of depots. These collections are frequently nested so they can be put on a one-dimensional density scale. As a consequence the higher level optimization procedure tends to amount to a search procedure over four parameters, the three channel rule parameters and the density of the regional depot set. Although an automatic search procedure over a four-dimensional grid could be superimposed on top of the lower level ALLOCRO-model, it is our experience that at this level an interactive approach is to be preferred, where the analysis of the output of the ALLOCRO-model for a given choice of higher level variables suggests which alternatives with respect to these variables should be considered next.

5.1 The ALLOCRO-model

Specification of the channel rule and the set of regional depots enables us to determine the set of direct destinations P and for each $p \in P$ the pair (δ_p, f_p) as defined in Section 4. This eliminates the need of considering the bottom level of the system in the remaining design optimization problem. The ALLOCRO-model determines the remaining three sets of variables; the set of DC's, the allocation of demand points in P to the DC's and the bulk flow variables. Two versions of the ALLOCRO-model will be presented, one in which the assignment of demand points to DC's is treated on an *individual* basis, and one in which *clusters* of demand points are allocated, where the clusters arise from a routing algorithm.

Before discussing these two alternatives we first present their common features. In both versions a further hierarchy is established among the three remaining sets of variables. The ALLOCRO-model generates a (locally) optimal set of DC's in an iterative fashion:

1. Given a current set of DC's an automatic or interactive mechanism generates a list of alternative sets, all of which differ from the current set in a single DC. E.g., an alternative set may add to or eliminate from the current set a DC (the greedy-add or greedy-drop

approach [CORNUEJOLS et.al 1977]). Another possibility is that an alternative set is obtained by replacing a DC within the current set by a DC outside of this set (the interchange heuristic or bump-and shift routine [KUEHN & HAMBURGER 1963]);

2. For the current set of DC's as well as for each alternative set a subroutine NETFLO determines the remaining two sets of variables, the territories of the DC's and the bulk flows of the commodities;
3. For the current set, a subroutine EVAL uses the output of NETFLO to compute the total system costs given the choices for the five sets of design variables and according to the principles described in Section 4;
4. The two versions of NETFLO use different procedures to obtain estimates for the total system costs of the alternative sets;
5. The alternative set with the lowest total system costs replaces the current set, provided that this replacement decreases the total costs. Otherwise the current set is a (local) optimum.

The two versions of ALLOCRO differ in the structure of the subroutine NETFLO and in the way in which estimates for the system costs of the alternative sets are obtained. We next turn to a discussion of the two versions of NETFLO.

5.2 NETFLO I

NETFLO I allocates demand points to the DC's on an *individual* basis and represents the traditional approach in location-allocation theory including the Geoffrion-Graves model [GEOFFRION et al. 1978]. In view of the occurrence of multiple delivery routes the cost of assigning an individual demand point to a DC is extremely hard if not impossible to evaluate. In fact the assignment costs are a nonlinear, "combinatorial" function of the assignment variables, defying the traditional literature. Crude linearizations have been suggested where the cost of assigning a demand point equals the cost of sending a special truck multiplied by some correction factor (cf. e.g. [MAIRS et al. 1978]). We discussed before that such linearizations may induce serious errors when used for estimating the transportation costs. NETFLO I uses the linearization merely to determine the territories of the DC's. Here the error is limited to the generation of suboptimal territory boundaries; potentially large deviations in the performance measures of the system cannot occur.

We next present the mathematical programming formulation of NETFLO I. For the sake of notational simplicity we assume that the throughput/ production costs in the plants and DC's are given by linear and separable functions in the throughput and production volumes. If this condition is not met we suggest as in [GEOFFRION & GRAVES 1971] to represent the original function by a piecewise linear approximation, and to introduce for each linear piece a fictitious facility with upper and lower limits on the throughput and production volumes. Now NETFLO I can be formulated as:

$$\text{(5.1)} \qquad \text{Minimize} \sum_{h,i,j,p} c_{hijp}\, u_{hijp} + \sum_j \phi_j \sum_p A_p v_{jp} + \sum_{j,p} \gamma_{jp}\, v_{jp}$$

$$\begin{aligned}
\text{s.t.} \quad SLO_{hi} \le \sum_{j,p} u_{hijp} &\le SUP_{hi} &&, \text{ for all } h,i \\
\sum_i u_{hijp} &= A_{hp} v_{jp} &&, \text{ for all } h,j,p \\
\sum_j v_{jp} &= 1 &&, \text{ for all } p \\
DLO_j \le \sum_p A_p v_{jp} &\le DUP_j &&, \text{ for all } j \\
u_{hijp} &\ge 0 &&, \text{ for all } h,i,j,p \\
v_{jp} &\in \{0,1\} &&, \text{ for all } j,p,
\end{aligned}$$

where

j runs over all open DC's,

p runs over all demand points P, each p characterized by a vector $(A_{1p},\ldots,A_{H_p})$, with

A_{hp} = the annual turnover of commodity h in point p (spu/yr),

$A_p = \sum_h A_{hp}$

u_{hijp} = the amount of commodity h, flowing from source i via DC_j to demand point p(spu/yr),

$$v_{jp} = \begin{cases} 1 & \text{if destination p is assigned to } DC_j, \\ 0 & \text{else} \end{cases}$$

$c_{hijp} = c^{bulk}_{hij}$ + (variable production or purchasing costs of h at source i /spu) + (variable production costs of commodity h at DC_j /spu)

ϕ_j = variable throughput costs at DC_j/spu

γ_{jp} = annual "linearized" transportation costs for deliveries from DC_j to demand point p.

A possible specification for γ_{jp} is:

$$\gamma_{jp} = \nu \frac{A_p}{Q} \cdot (t_{jp}\, vt + d_{jp}\, vm), \tag{5.2}$$

with

t_{jp} = the travel time from DC_j to demand point p,

d_{jp} = the distance from DC_j to demand point p,

ν = a correction factor for direct deliveries.

In our case study we determined the correction factor ν as follows. We selected a recent pilot year and set ν such that NETFLO I's estimate of the ratio of the number of miles for direct DC-deliveries and the number of miles driven on bulk transportation routes, this ratio multiplied by ν, equals the actually realized ratio of these quantities in the chosen pilot year. Different correction factors have e.g. been used by Mairs [MAIRS et al. 1978].

5.3. NETFLO II

In NETFLO II *clusters* of demand points are allocated to the open DC's. These clusters correspond with routes in which the links with the DC have been replaced by a link from the last demand point on the route to the first one. As an example (Figure 2), the cluster (1,...,5) has been obtained from the circuit (0,1,...,5,0) by deleting the links (0,1) and (5,0) and adding link (5,1).

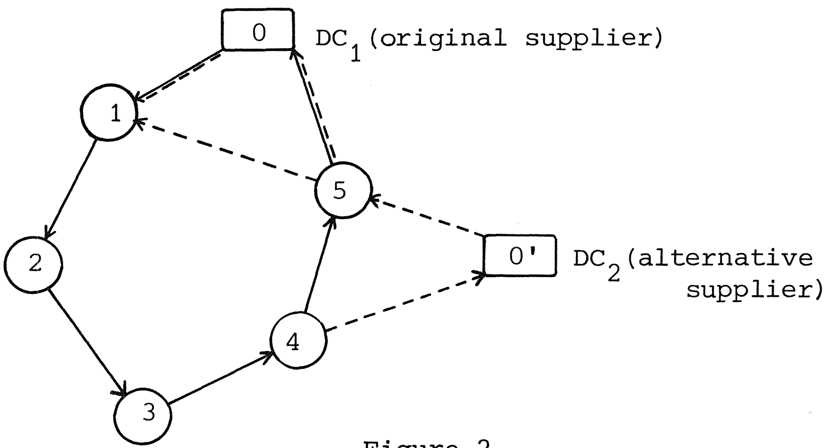

Figure 2

Note that a cluster r of m demand points is specified by an *ordered m-tuple* of locations $(p_1,...,p_m)$.

We mentioned before that at each iteration of the ALLOCRO-model the costs associated with the current set of DC's are evaluated by the subroutine EVAL which determines routes for each of the DC-territories for each day of the standard week. For a given current set and a given standard week, consider the set of routes generated by EVAL and let R be the corresponding set of clusters. In the process of the ALLOCRO II-algorithm the set R changes at times, namely whenever a new set of DC's becomes "current" and the subsequent application of EVAL generates a new set of routes. NETFLO II, applied to some set J^0 of open DC's uses the "current" collection R and allocates each cluster in R to one DC in J^0. To initialize the set R a different route-clustering procedure or an altogether different version of NETFLO will have to be used.

To facilitate the subsequent presentation we explain the algorithm for the case where all demand points want daily service ($\hat{f}_p=1, p\in P$). Extensions to the general multi-frequency case have been developed and will be described in a forthcoming paper.

For a given cluster r we define the *extension* of r to DC_j as the cost minimal route which starts at DC_j, has one of the points in r as its first delivery point, visits the remaining points according to the permutation of r and returns to the DC. E.g. in Figure 2, the extension of cluster $\{1,\ldots,5\}$ to DC_2 is the route $(0',1,2,3,4,0')$.

Let $\tilde{c}_{pq}$ resp. $\tilde{\gamma}_r$ be the cost of traversing link (p,q) resp. route r. Defining $\tilde{\gamma}_{jr}$ and $\tilde{t}_{jr}$ as the cost resp. the duration of the extension of cluster r to DC_j we have:

$$\tilde{\gamma}_{jr} = \tilde{\gamma}_r + \min_{1\le\ell\le m} \{ \tilde{c}_{jp_\ell} + \tilde{c}_{p_{\ell-1}j} - \tilde{c}_{p_{\ell-1}p_\ell} \},$$

where the index of p is assumed to be taken modulo m. If $\tilde{t}_{jr} \le T$, this number can be taken as the assignment costs γ_{jr} of r to DC_j, since in this case the extension of r to DC_j represents a feasible route. Otherwise, we divide tour $(p_1,\ldots,p_m)$ into two Hamilton paths, in such a way that the extensions of the two corresponding circuits to DC_j are feasible and the sum of the costs of those extensions is minimal:

$$(5.3) \qquad \gamma_{jr} = \begin{cases} \tilde{\gamma}_{jr} & \text{, if } \tilde{t}_{jr} \le T \\ \min\{\tilde{\gamma}_{jr_1}+\tilde{\gamma}_{jr_2} \mid r_1 = (p_\ell,\ldots,p_q), r_2 = (p_{q+1},\ldots,p_{\ell-1}), \\ \quad 1 \le \ell < q \le m,\ t_{jr_1} \le T,\ t_{jr_2} \le T\}, & \text{otherwise,} \end{cases}$$

with the understanding that the minimum of an empty set equals $+\infty$.

Note that the computation of γ_{jr} requires either $O(m)$ or $O(m^3)$ operations. Even though m is usually less than say 15, this may be excessive. In that case it may be preferable to consider only a subset of the possible divisions of cluster r, e.g. into Hamilton paths with a (nearly) equal number of points, which reduces the computational effort to $O(m^2)$.

Defining A_{hr} as the annual turnover of commodity h summed up for all demand points in cluster r, and similarly A_r as $\Sigma_h A_{hr}$, we are now ready to give the mathematical programming formulation of NETFLO II. In fact, the formulation of NETFLO II is identical to that of NETFLO I with two exceptions:

- the index p does not run over the set of demand points P, but over the set of clusters R;
- the coefficients γ do not represent "linearized" transportation costs computed according to (5.2) but cluster assignment costs computed according to (5.3).

5.4 Conclusion

The two versions of NETFLO represent special cases of the Geoffrion-Graves model [GEOFFRION & GRAVES 1971] and as such they can be solved via a decomposition method with a master program involving the allocation variables and a transshipment problem for each commodity as a subroutine. One-commodity models reduce to simple transshipment problems when the condition that each demand point is supplied from a single DC is relaxed. Counting the number of basic variables in this problem one concludes that the number of demand points receiving deliveries from more than one DC is at most equal to the number of sources plus the number of open DC's, which is usually a very small fraction of the number of demand points. Hence in one-commodity models the transshipment relaxation leads to feasible or near-feasible solutions with a feasible and optimal solution guaranteed via embedding in a small branch-and-bound procedure. In fact this device can sometimes be extended to multicommodity models with the special property that only one commodity allows a choice between multiple sources to supply the DC's. This possibility was exploited in our case-study (cf. Section 6).

We summarize our presentation of the two versions of ALLOCRO in the two flowcharts of Figure 3. The differences of the two versions can be described as follows.

1. In ALLOCRO I the subroutine EVAL is used to determine the total systems costs for every alternative set rather than being applied only to the sets which in the course of the algorithm become "current". As explained above, NETFLO I serves to provide the DC-territories and the bulk flows as input variables for the EVAL subroutine.
2. In ALLOCRO II the subroutine EVAL is merely invoked for those sets of DC's which in the course of the algorithm become designated as "current". For every alternative set, the optimal value of NETFLO II is used as an estimate for the costs of bulk transportation, variable throughput costs in the DC's and the costs of direct deliveries. This is justified since the third component of the objective function of NETFLO II (cf. 5.1) represents the "exact" routing costs for a (suboptimal) solution to the routing problem. The remaining components in the total systems cost function are next added along the lines described in Section 3 at neglectible additional effort.

We conclude this section with the following remarks. NETFLO I tends to slightly overestimate the real costs of direct deliveries in view of possible gains obtainable by rerouting the clusters that have been assigned to a DC-territory. One such possibility occurs when several clusters have been cut into two subclusters in order to be assigned to a particular DC (cf. Section 5.3). In this case gains may be obtained by combining pairs of subclusters into clusters. A possible extension of NETFLO II would therefore add a procedure in which subclusters allocated to the same DC-territory would be matched via a weighted matching algorithm.

Finally we mention the following two advantages of ALLOCRO II as compared to ALLOCRO I. First, several alternative sets can be considered between subsequent uses of the expensive EVAL routine. Moreover the allocation of clusters of demand points provides a natural aggregation procedure. In fact the size of the problem is reduced by a factor ε, where ε is defined (see Appendix 2) as the average number of deliveries on a route.

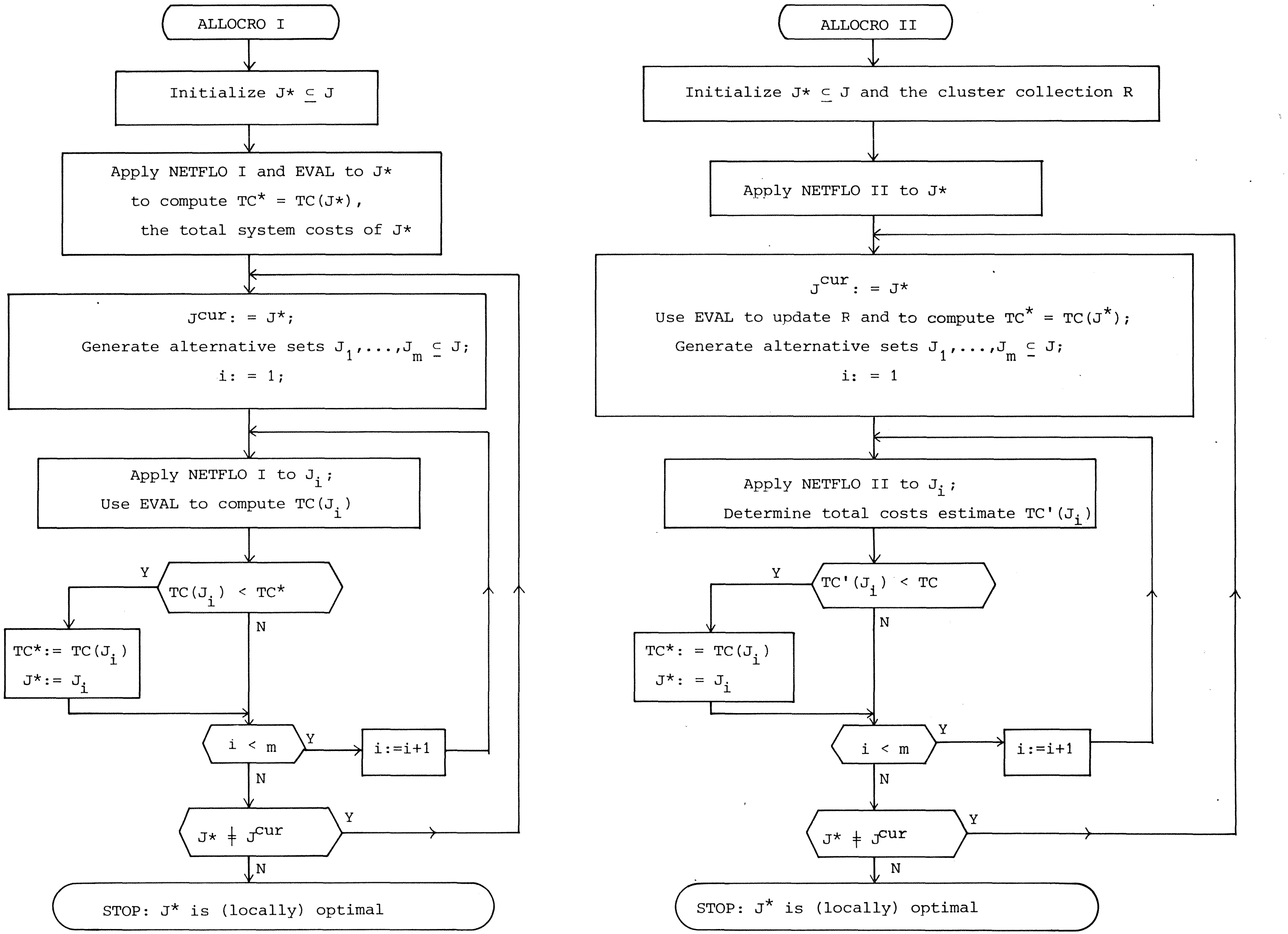
ALLOCRO I
Initialize J* ⊆ J
Apply NETFLO I and EVAL to J*
to compute TC* = TC(J*),
the total system costs of J*
Jcur: = J*;
Generate alternative sets J1,...,Jm ⊆ J;
i: = 1;
Apply NETFLO I to Ji;
Use EVAL to compute TC(Ji)
TC(Ji) < TC*
Y
N
TC*:= TC(Ji)
J*:= Ji
i < m
Y
N
i:=i+1
J* ≠ Jcur
Y
N
STOP: J* is (locally) optimal
ALLOCRO II
Initialize J* ⊆ J and the cluster collection R
Apply NETFLO II to J*
Jcur: = J*
Use EVAL to update R and to compute TC* = TC(J*);
Generate alternative sets J1,...,Jm ⊆ J;
i: = 1
Apply NETFLO II to Ji;
Determine total costs estimate TC'(Ji)
TC'(Ji) < TC
Y
N
TC*: = TC(Ji)
J*: = Ji
i < m
Y
N
i:=i+1
J* ≠ Jcur
Y
N
STOP: J* is (locally) optimal

Figure 3

6. SOLUTION OF A LARGE PRACTICAL PROBLEM

AGA Gas B.V., the Dutch daughter of AGA Sweden and Air Liquide in France, produces and distributes a variety of industrial gases. All but a small percentage of the customers receive their deliveries in gas cylinders and correspondingly our study was confined to the cylinder distribution process. Although the product line consists of several scores of items including various mixtures of gases it suffices to distinguish between three product categories: oxygen, acetylene and other gases. This 3-partition is justified by the fact that oxygen and acetylene represent respectively 50% and 35% of the turnover as measured in number of cylinders; moreover only with respect to oxygen and acetylene the distribution process originates from multiple sources, either external suppliers or company owned plants.

The distribution network of the company satisfies the description given in Section 2 (see Figure 1). In particular, the multicommodity aspect of the system is irrelevant for the part of the distribution process between the DC's and the final destinations. This is due to the fact that the various gases are pumped into cylinders of a standard size; this pumping process either takes place at the source or at the latest at the DC. In addition each customer and depot receives all products from the same DC.

The internal production process of oxygen and acetylene is characterized by two stages, one in which the half products liquid oxygen and carbide are produced, and a so-called filling stage in which the half products are transformed into gaseous oxygen resp. acetylene and subsequently pumped into cylinders. For the two products concerned we undertook a preliminary study to point out which of the potential DC's should receive deliveries as half-products rather than as finished article. In other words, this study determined the sets $S(h,j)$ as defined in Section 2, using an analysis which traded economies of scale in the production area against differences in transportation costs between gas cylinders and the substantially less voluminous half products.

With respect to the third product category, the dispatching of all other cylinders originates from a single company owned plant, where some of the products are produced and other purchased.

A preliminary operational accounting study established the cost functions $\Theta(\cdot)$ and $\Phi(\cdot)$ by representing each of the labour and capital cost components separately as a function of the throughput and production volumes.

In the existing situation the distribution network of the company included four DC's and approximately 100 depots. Management presented us a list of ±20 potential DC locations. With respect to the depots, depotsets were created with a more or less uniform dispersion over the Netherlands and with cardinality varying between 45 and 200.

It is worthwhile to note that in AGA's case the depots are no company owned facilities but independent agents, performing the last stage of the distribution process for a fixed commission per cylinder. Each depotholder covers his own handling and transportation costs from his revenue of commission and transportation fees he is allowed to charge to his customers. In spite of the economic independence of the depots it was AGA's privilege to determine the channel choice problem, including the allocation of customers to depots. In other words AGA has the option of shifting customers from and to the depots within the hazy constraint of preserving sufficient income for the depotholder to keep his agency profitable while on the other hand not overburdening his handling capacity.

Because no relationship between the depot's commission fee and the customer composition of the depot's territory was available or predictable, we decided to strive after a design of the distribution network that minimizes the total systems costs rather than AGA's own costs. In other words we kept the transportation cost between depots and indirect customers as part of the objective function and treated the commission as the renumeration of the depotholder's marketing and administrative efforts solely. To represent the depotholder's transportation costs we used estimates of the mileage fees that external local carriers would charge. This seemed the only reasonable estimation procedure in a situation where the biggest possible heterogeneity prevailed with respect to the transportation structure of the depots, due to:

- the percentage of customers in the depot's territory that chose to pick-up their cylinders themselves;
- the vehicle fleet available;
- the extent to which the depotholder combined other retail or wholesale

activities with the AGA deliveries.

We next discuss the composition of the customer set. The analysis could be restricted to ±3000 account customers with an annual turnover of more than 25 cylinders. An additional 17.000 customers were receiving incidental deliveries, mostly on a strictly cash-on-delivery basis. The contribution of this group to the total turnover in cylinders was less than 12%. No channel choice or allocation problem arose for this group of customers; for each set of DC's and depots under consideration, its turnover was allocated to individual DC's and depots in proportion to the allocated turnover for account customers. For the remaining 3000 customers an optional channel choice rule was found in combination with a solution to the other network design problems.

Finally we note that the above discussed multicommodity aspect of the system was further restricted by the fact that each DC had a single predetermined source both for oxygen and for the third product category of other gases. As mentioned in the previous section this special feature could be exploited to formulate the NETFLO I problem as a transshipment problem. Neglecting the capacity constraints on the DC's and denoting acetylene, the only commodity without predetermined sources, as commodity 1, we are faced with the following problem:

$$\text{minimize} \quad \sum_{i,j} c_{1ij} u_j + \sum_{h \neq 1} \sum_j c_{h \cdot j} \sum_p \frac{A_{hp}}{A_{1p}} v_{jp} + \sum_j \phi_j \sum_p \frac{A_p}{A_{1p}} v_{jp} + \sum_{j,p} \frac{\gamma_{jp}}{A_{1p}} v_{jp}$$

$$\text{s.t.} \quad SLO_{1i} \leq \sum_j u_{ij} \leq SUP_{1i}, \qquad \text{for all } i$$

$$\sum_i u_{ij} = \sum_p v_{jp}, \qquad \text{for all } j$$

$$\sum_j v_{jp} = A_{1p}, \qquad \text{for all } p$$

$$u_{ij}, v_{jp} \geq 0, \qquad \text{for all } i,j,p$$

where the notation is the same as for NETFLO I (Section 5), except that

u_{ij} = the amount of acetylene cylinders, flowing from source i to DC_j (cyl/yr),

v_{jp} = the amount of acetylene cylinders, flowing from DC_j to demand point p (cyl/yr),

c_{1ij} = costs of 1 acetylene cylinder at DC_j originated from source i (costs of production/purchasing and bulk transpotation)

$c_{h \cdot j}$ = costs of 1 cylinder of commodity h at DC_j.

By counting the number of basic variables and constraints in this transshipment problem, one can easily see that the number of demand points receiving split deliveries is at most $|I| + |J^0|$ in any basic solution of this problem, a neglectible fraction of $|P|$; in all our runs not a single split delivery occurred.

We conclude this section by presenting some results.

Figure 3a. Territories for a 3 - DC - set

Figure 3 shows the territories associated with two of the DC- sets under consideration, where the DC is printed for each demand point.

Figure 3b. Territories for a 4 - DC - set

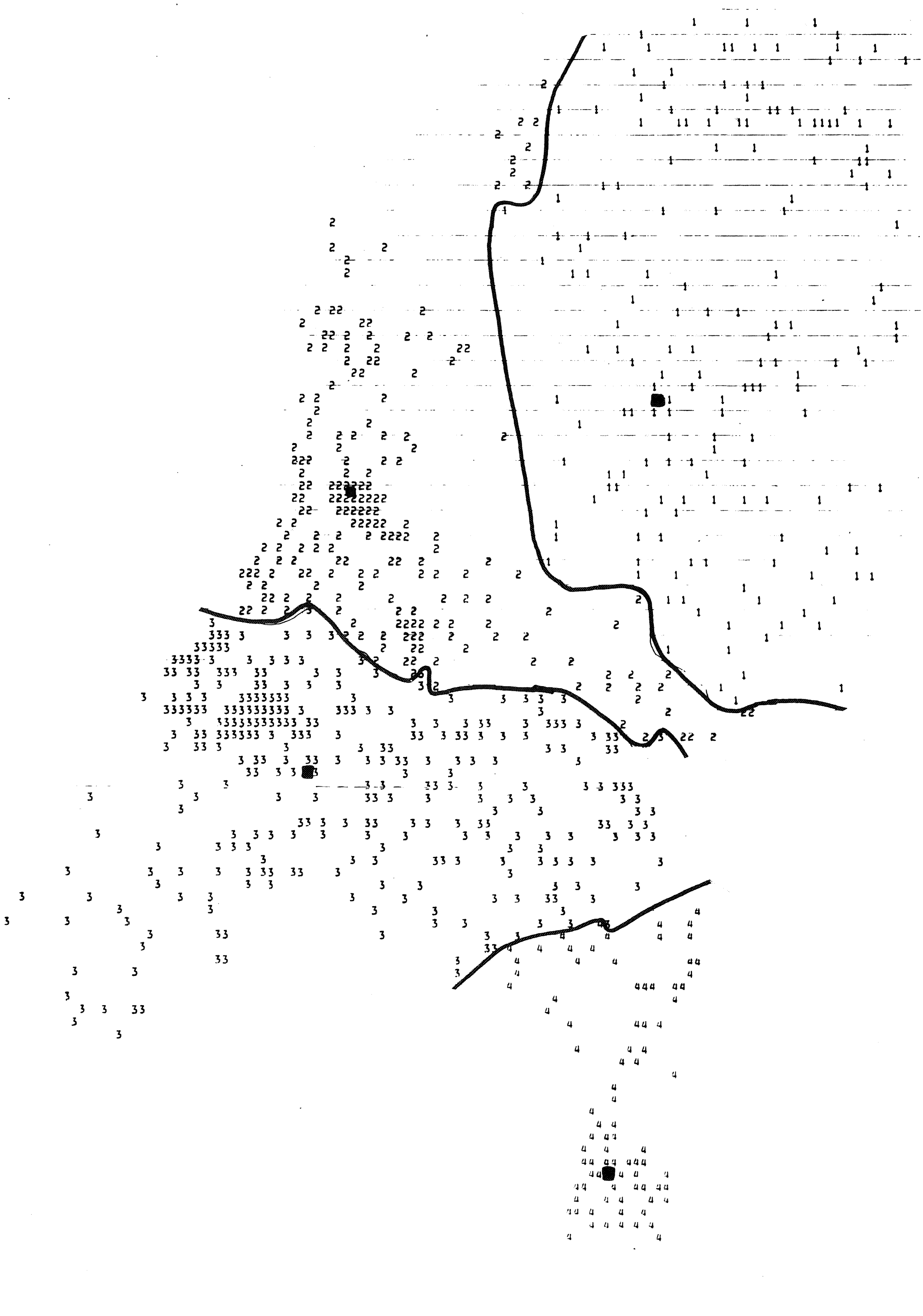

Figure 4. Map of the separation degree (scale 0 - 10)

Figure 4 exhibits the distribution of the separation degree over the Netherlands, where this degree is mapped on a scale from 0 to 10.

Table 1. Influence of the channel rule on the ratio direct/indirect cylinders

Channel rule δ_-	δ^+	σ^+	Indirect Cylinders (*1000)	Direct Cylinders (*1000)	% Indirect Cylinders
4	12	50	330	630	34
8	8	-	365	595	38
8	16	75	415	545	43
4	12	25	435	525	45
8	24	50	495	465	52
16	16	-	575	385	60
1977			580	380	61
8	24	25	640	320	67
32	32	-	800	160	83

Table 1 shows that the impact different channel rules have on the volumes allocated to the direct and indirect channel.

Table 2. Survey of variable cost components for a given DC- and depot-set.

Channel rule			costs of				total of
δ_-	δ^+	σ^+	(1) direct deliveries	(2) indirect deliveries	(3) depot's commission	(4) inventory at DC's and depot's	(1) - (4)
4	12	50	61	17	19	11	108
8	8	-	53	16	20	11	100
8	16	75	48	21	23	12	103
4	12	25	47	22	24	12	105
8	24	50	45	26	26	13	110
16	16	-	41	29	30	13	113
1977			44	28	31	13	115
8	24	25	39	32	33	13	117
32	32	-	36	40	40	14	130

Table 2 shows for a given DC- and depot-set how the various cost components fluctuate as a function of the channel rule.

Table 3. Output of vehicle routing module in EVAL

SUMMARY OF DELIVERIES

DAY	DELIVERIES	CYLINDERS
1	52	1281
2	49	959
3	52	1118
4	48	1138
5	55	1234

ROUTES FOR DAY 1

ROUTE	TRUCK	LENGTH	TIME[1]	DEL.[2]	CYL.[3]	ROUTE:[4]
1	1	272	507	6	104	430 11 297 7 9
2	3	186	465	6	157	402 314 315 316 318 14
3	4	148	426	5	153	775 458 19 311 25
4	2	148	495	7	132	31 485 24 490 486 21 396
5	5	26	495	9	157	787 844 726 748 724 746 725 841 797
6	7	13	179	2	160	33 774
7	8	6	238	6	151	773 798 772 789 852 857
8	6	75	400	7	153	718 701 700 702 704 706 699
9	7	65	301	4	114	36 903 979 42
DAY TOTAL	8	939	3416	52	1281	
AV./TRUCK		117	427	6	160	
AV./ROUTE		104	379	5	142	

ROUTES FOR DAY 2

ROUTE	TRUCK	LENGTH	TIME	DEL.	CYL.	ROUTE:
1	1	248	510	6	68	1639 2166 2169 2175 78 2174
2	4	238	492	6	101	2154 2156 2159 2163 2165 73
3	5	240	446	3	73	1919 65 2002
4	2	247	509	7	140	1015 39 47 51 1301 1220 53
5	3	151	500	10	160	941 966 1059 1454 1557 1517 1486 1441 58 37
6	6	158	404	6	152	946 1020 61 62 688 760
7	7	69	384	7	156	792 34 700 699 842 848 843
8	8	12	185	4	109	773 793 775 772
DAY TOTAL	8	1363	3430	49	959	
AV./TRUCK		170	428	6	119	
AV./ROUTE		170	428	6	119	

ROUTES FOR DAY 3

ROUTE	TRUCK	LENGTH	TIME	DEL.	CYL.	ROUTE:
1	3	261	449	5	112	11 12 7 209 259
2	1	188	502	7	154	451 313 315 316 324 325 14
3	7	81	303	5	156	800 775 501 502 25
4	5	72	381	6	159	799 699 702 31 686 700
5	6	25	356	8	159	772 787 773 851 764 763 838 989
6	2	140	481	8	153	1147 1249 59 1358 50 1319 52 1235
7	4	89	433	8	119	975 841 1058 44 1129 1112 900 835
8	8	90	285	5	106	55 1407 1393 1435 1392
DAY TOTAL	8	946	3190	52	1118	
AV./TRUCK		118	398	6	139	
AV./ROUTE		118	398	6	139	

ROUTES FOR DAY 4

ROUTE	TRUCK	LENGTH	TIME	DEL.	CYL.	ROUTE:
1	1	250	508	6	42	2164 2165 2168 2179 78 2173
2	4	228	491	6	126	65 2155 2110 2158 2160 73
3	3	277	499	6	122	56 53 1661 1519 58 51
4	5	109	391	8	153	1015 1557 1654 62 1517 1498 1454 953
5	6	112	342	6	153	881 988 1023 49 1300 39
6	2	175	506	7	153	699 660 847 1020 1048 1022 61
7	7	59	341	6	149	775 791 700 685 927 790
8	6	10	124	2	80	773 772
9	8	184	336	1	160	311
DAY TOTAL	8	1404	3538	48	1138	
AV./TRUCK		175	442	6	142	
AV./ROUTE		156	393	5	126	

ROUTES FOR DAY 5

ROUTE	TRUCK	LENGTH	TIME	DEL.	CYL.	ROUTE:
1	3	126	462	9	152	942 970 1081 50 1464 1478 59 1463 1411
2	2	112	493	9	115	1386 1415 1416 1453 1413 52 1319 1320 1321
3	1	101	493	11	132	1143 1190 1393 55 1511 1484 57 1405 1064 1024 950
4	5	126	447	7	154	949 42 44 1130 1259 841 844
5	4	124	451	8	151	787 813 789 702 36 47 40 37
6	7	61	258	3	154	661 33 775
7	6	108	360	5	136	773 694 503 700 699
8	6	6	140	1	160	771
9	8	6	84	1	48	772
10	7	104	210	1	32	19
DAY TOTAL	8	874	3398	55	1234	
AV./TRUCK		109	424	6	154	
AV./ROUTE		87	339	5	123	

TRANSPORTATION REPORT

	WEEK TOTAL	AVERAGE/TRUCK	AVERAGE/ROUTE
TRUCKS	8		
ROUTES	44		
KILOMETERS DRIVEN	5526	690	125
TIME DRIVEN	16972	2121	385
DELIVERY POINTS	256	32	5
CYLINDERS	5730	716	130
WORK LOAD[5]		89	
BUSY TIME (PROC)		83	
KILOMETERS DRIVEN/CYLINDER	0.96		

1) time in minutes; maximum 510 min.

2) number of delivery points

3) number of cylinders delivered

4) number of consecutive delivery points on route

5) standard = 100, corresponding with 1 full truck of 160 cylinders.

Finally Table 3 gives the output of the vehicle routing module in EVAL with respect to one of the DC's under consideration.

APPENDIX 1

THE COMPUTATION OF THE SEPARATION DEGREE

This appendix presents a mathematical programming formulation as well as heuristic and exact solution methods for the tour selection problem which underlies the computation of the separation degree on a particular standard day for a given customer which we will denote as the *seed customer*.

First however a single remark should be made regarding the generation of standard days which as we recall is based upon the assumption that the occurrence of individual customers represents a series of statistically independent events. In practice vehicle routing algorithms tend to assign neighbouring customers to the same set of days, thus exploiting their geographical proximity. This implies that the conditional probability of a demand point in the neighbourhood of the seed customer appearing on a typical day on which the latter is to be served exceeds the unconditional probability, the difference between the conditional and the unconditional probability tending to decrease as the distance to the seed customer increases. This positive correlation between the occurrence of neighbouring demand points on a standard day should ideally be accounted for. The independence assumption used by us thus is prone to slightly overestimate the "truely desired" value, especially in areas with a high density of low frequency customers.

Let 0 denote the seed customer and let 1,2,...,n represent the remaining customers in the standard day (NB: substantial computational savings can be attained by restricting the set to those customers that lie within an a priori determined radius from the seed customer). The separation degree equals the value of the "tour selection" problem which can be formulated as follows:

$$\text{(P)} \quad \text{minimize} \quad \sum_{i=0}^{n} \sum_{j=0}^{n} t_{ij}\omega_{ij} \qquad (1)$$

$$\text{s.t.} \quad \sum_{j=0}^{n} \omega_{ij} = 1 \qquad i = 1,\dots,n \qquad (2)$$

$$\sum_{i=0}^{n} \omega_{ij} = 1 \qquad j = 1,\dots,n \qquad (3)$$

$$\sum_{i=0}^{n} \sum_{j \neq i} \delta_j \, \omega_{ij} \geq Q \tag{4}$$

$$\sum_{\substack{\{i,j\} \subseteq V \\ i \neq j}} \omega_{ij} \leq |V| - 1, \quad V \subseteq \{1,\ldots,n\} \quad 2 \leq |V| \leq n-1 \tag{5}$$

$$\omega_{ij} = 0,1 \qquad i,j = 0,\ldots,n, \tag{6}$$

where

$$\omega_{ij} = \begin{cases} 1 & \text{if the tour visits j after i} \\ 0 & \text{otherwise.} \end{cases}$$

t_{ij} = variable cost of direct travel from customer i to customer j. Usually the shortest route in distance or time on a road network is used, although a cost weighted combination is preferable. Travel times may include loading/unloading at the customers, given e.g. as a function of the dropsize. We set $t_{ii} = t_{i0} = 0$, for all i.

δ_i = dropsize of customer i

Q = capacity of a vehicle (in spu).

Problem (P) is a linear assignment problem with additional constraints. Constraint (4) ensures that at least the truck capacity is delivered on the tour. The constraint set (5) denotes the classical subtour eliminating constraints of the traveling salesman problem for the remaining demand points 1,...,n. Constraints (5) together with the assignment constraints (2), (3) and (6) ensure that a solution of (P) consists of a (sub)tour through the seed customer and loops on the demand points not in this (sub)tour. Because of the dummy links (i,0) and the capacity constraint (4), this subtour really is a path originating at the seed customer with a load at least equal to the vehicle capacity.

The tour selection problem has two aspects: a selection aspect and a routing one. One can easily formulate the subset sum problem [GAREY & JOHNSON 1979] as a tour selection problem. On the other hand a traveling salesman problem on n points with distance matrix (d_{ij}) can be solved by the tour selection problem with

$$Q = n+1$$

$$\delta_i = 1 \qquad ,i = 0,\dots,n$$

$$t_{ij} = \begin{cases} 0 & ,j = 0 \text{ or } i = j \\ d_{1j} & ,i = 0,\ j \geq 1 \\ \Sigma_{i,j}\, d_{ij} & ,i = 1,\ j \geq 2 \\ d_{ij} & ,\text{else} \end{cases}$$

Hence the tour selection problem with no restriction on the value of Q is a generalization of the travelling salesman problem and NP-complete in the strong sense.

In practice however most times no more than 20 customers are needed to construct the optimal tour and frequently even less than 10. This shows that for practical problems most of the computational effort has to be directed towards the selection of the right subset of demand points rather than towards the routing of this subset.

We conclude this appendix with a number of solution methods for the tour selection problem.

Branch-and-bound methods

A first method that comes to mind is the relaxation of (P) where the subtour eliminating constraints (5) are removed. The relaxed problem is a linear assignment problem with (4) as one extra constraint. The optimal solution of this problem [KLINGMAN & RUSSELL 1975] yields a lowerbound on the separation degree of the seed customer; it is feasible and hence optimal for (P) if all associated tours which are not incident to the seed customer are loops. Otherwise a branch-and-bound tree search is needed where the branching rules can be taken as in the classical branch-and-bound methods for the traveling salesman problem (cf. Section 9.2.2 of [LENSTRA 1977]). We note that the size of the search tree may be at least of the same order as for the simple traveling salesman problem; therefore this exact solution method is intractable for large size problems other than for the purposes of obtaining a lowerbound on the separation degree.

Methods based on Benders Decomposition

This method resembles the vehicle routing algorithm of Fisher and Jaikumar [FISHER & JAIKUMAR 1978]. The algorithm is based on a different formulation of the tour selection problem which lends itself to a decomposition into a knapsack masterproblem - with additional constraints from Benders cuts at later iterations -, which determines the customers on the route, and a traveling salesman subproblem determining the minimal Hamiltonian path through this set. Although in principle an exact method a suboptimal variant of this algorithm with termination after a limited number of iterations seems more promising. Heuristics for the choice of an initial objective function in the knapsack variables are essential for a successful implementation of the algorithm.

Heuristic Route-Building and Improvement Algorithms

A third category of algorithms guarantees local rather than global optimality even when pursued until convergence. For this type of algorithm two phases can be distinguished: a first phase in which an initial route is constructed and a second one in which improvements are achieved through interchanges.

The improvements procedures resemble the 2-opt methods that have proven to be highly successful for the traveling salesman problem [LIN 1965; LIN & KERNIGHAN 1973]: in each iteration one considers feasibility preserving replacements of p customers currently belonging to the route by r customers outside of the route. The case $p \neq r$ has to be used since the number of customers on the route is unknown a priori.

For the initial route-building part we mention the following greedy type class of algorithms, which is based on the knapsack relaxation of (P):

Greedy Route-Building Algorithm K

1. At the m-th step, a path $(\ell_0 = 0, \ell_1, \ldots, \ell_m)$ has been constructed. If $\Sigma_{k=0}^{m} \delta_{\ell_k} \geq Q$, the path terminates. Otherwise a new customer $\ell_{m+1} = j$ is added for which

$$\frac{\tau_j}{\delta_j} = \min_{\ell \notin \{\ell_1, \ldots, \ell_m\}} \frac{\tau_\ell}{\delta_\ell},$$

with

$$\tau_\ell = \min_{m-K \leq k \leq m} t_{\ell_k \ell}.$$

2. An exact or heuristic algorithm for the traveling salesman problem constructs a Hamiltonian path through the generated set of customers.

APPENDIX 2
AN INITIAL ASSIGNMENT OF DEMAND POINTS TO THE DAYS OF THE STANDARD PERIOD

This appendix describes the heuristic method that was used in the case study to create an initial assignment of demand points to the days of the standard period. Although this method was tailored to the specific properties of our problem, we believe that the main principles are of general use.

In our problem we have $W = 5$, i.e. a standard period of 5 working days, and $\hat{f}_p = 1,2,3$ or 5. Customers with $\hat{f}_p = 4$ are missing because the company's distribution managers when planning basic periodic delivery schedules did not distinguish a separate class of destinations with frequency 4. Hence in the creation of the standard week values of f_p between 3 and 5 were rounded up to 5 with probability $(f_p-3)/2$ and down to 3 with the complementary probability. The spacing conditions required destinations with $\hat{f}_p = 3$ to be placed on Monday, Wednesday and Friday, and destinations with $\hat{f}_p = 2$ to be placed on any pair of nonconsecutive days.

Our heuristic tries to balance the routing costs on the W days, as justified in Section 4. To meet this objective, we use an approximation due to Eilon et al. [EILON et al. 1971] for problems where the locations of the customers are homogeneously distributed within a square of side-length a. For a set U of points to be visited on the same day, let $d_{pp'}$ be the travel time between points p and p' in U; let $D_0(U)$ be the total duration of the routes, $D_r(U)$ the sum of the direct "radial" travel times d_{0p} between the DC and the destinations in U, and $\varepsilon(U)$ the average number of deliveries per route. The following approximate relationship was found to hold:

$$D_0(U) \sim \left[\frac{AD_r(U)}{\varepsilon(U)} + B\sqrt{a}\,\sqrt{D_r(U)}\right] q(s).$$

Here A and B are regression coefficients ($A \sim 1.75$ and $B \sim 1.05$) with remarkable robustness with respect to the location of the DC and the distribution of the dropsizes; s represents the variational coefficient of the dropsizes and $q(\cdot)$ is a given monotonically increasing function of s.

This approximation formula has proven to be relatively accurate in quite a number of cases, especially when a good estimate for $\varepsilon(U)$ can be found. This becomes easier when the bottleneck of a route is due to a single constraint, e.g. the volume capacity of the vehicle or the maximum permitted time-length, and when estimates exist for the average dropsize or the average travel time between two consecutive delivery points. The linear term in the above formula dominates when the number of delivery points $n \to \infty$ and $\varepsilon(U)$ is small, with $\varepsilon(U) = 1$ or individual service as an extreme case. In the other extreme where $\varepsilon \to c.n$ as $n \to \infty$, with $0 < c \leq 1$, the VRP solution approaches asymptotically the length of 1/c traveling salesman tours. In this case $D_0(U)$ behaves asymptotically as a multiple of the square root of $D_r(U)$, confirming recent asymptotic results by Stein [STEIN 1978].

For the parameter a we choose

$$\hat{a}(U) = \sum_{\substack{\{p,p'\}\in U \\ p\neq p'}} d_{pp'} \Big/ \left(.52\binom{n}{2}\right),$$

since $.52\,\hat{a}(U)$ is an unbiased and asymptotically consistent estimator of

$$\frac{1}{a^4}\int_0^a\int_0^a\int_0^a\int_0^a \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2}\, dx_1 dx_2 dy_1 dy_2 \sim .52\, a,$$

the expected value of the distance between two points in a square with side-length a, in which the points are uniformly distributed.

Let $\hat{P}$ be the set of delivery points in the standard week in which delivery point p is represented $\hat{f}_p$ times. As justified above, we opt for equalization of the workload on the various days; furthermore we need the assumption that both the daily sum of radial travel times and the daily sum of loading times will be equally distributed over the standard period and hence be equal to $D_r(\hat{P})/5$ and $\Lambda(\hat{P})/5$. We thus put

$$D_0^*(\hat{P}) = \left\{\frac{A}{\varepsilon(\hat{P})}\frac{D_r(\hat{P})}{5} + B\sqrt{\hat{a}(\hat{P}).\frac{D_r(\hat{P})}{5}}\right\}q(s) + \frac{\Lambda(\hat{P})}{5}$$

as the target value of the total duration of the routes on a single day loading times included.

The assignment of the customers to days of the week is performed on a sequential basis in descending order of frequency of delivery. First the points with frequency 5 or 3 are assigned since for these only one day combination is permitted. Next we deal with the demand points with $\hat{f}_p = 2$ and finally the once-a-week destinations are assigned. Let P_w, w = 1,...,5, represent the set of demand points already assigned to day w. A day w is *open for demand point* p if $D_0(P_w+p) + \Lambda(P_w+p) \leq D_0^*(\hat{P})$. Within each frequency class we assign at each iteration the demand point that has the best *proximity value* (to be specified below) for one of the permitted day combinations. A day combination is permitted for a particular demand point if it satisfies the spacing conditions for this point and if all days of the combination are open for this point. The proximity values for a particular demand point and day combination depend on the sets P_w, w = 1,...,5, and are as a consequence dynamically adapted. They are computed as a weighted sum of the proximity values for each day in the combination. Before discussing some of the possible choices for these values we point out that this nethod satisfies the principle of creating a *geographic cohesiveness* among destinations assigned to the same day. A full justification for this principle can only be given in case the fleet of vehicles consists of a single truck. The success of an algorithm based upon this principle is as surprising as the success of say savings criteria in TSP- and VRP-problems which are based on the desirability of combining two points into a tour of two delivery nodes.

Possible choices for the proximity value π_{pw} of demand point p to day w include:

1. The Clarke & Wright savings with respect to one of the customers in P_w:

$$\pi_{pw}^{(1)} = \min_{p' \in P_w} \{d_{0p} - d_{pp'} + d_{p'0} \mid \delta_p + \delta_{p'} \leq Q\}$$

   Note that only pairs of customers with a combined demand not exceeding the vehicle capacity are considered;

2. Assume a tree has been spanned on P_w rooted at the DC. A second alternative for the proximity value is the saving incurred when connecting p to the tree:

$$\pi_{pw}^{(2)} = d_{0p} - \min_{p' \in P_w \cup \{0\}} \{d_{pp'} \mid \delta_p + \delta_{p'} \leq Q\}$$

   If p is assigned to day w, the tree is extended with the link that realizes $\pi_{pw}^{(2)}$;

3. Assume a Hamiltonian circuit has been generated through the points in $P_w \cup \{0\}$. A third alternative are the minimial insertion costs when inserting p in between two consecutive delivery points on the route. If p is assigned to day w it is inserted in between the minimizing pair of consecutive points.

The attraction power of a day depends on the number of customers already assigned to it. In order to prevent snowball effects two devices were found to be useful:

- the assignment procedure is executed in two phases. In the first phase the upperbound for determining whether a day is open for a customer is set equal to 1/2 $D_0^*(\hat{P})$; in the second phase the full upperbound $D_0^*(\hat{P})$ is used;
- instead of using the unweighted sum of the proximity values over the days in the day combination, we may weigh them with a factor which is monotonically increasing in $D_0^*(\hat{P}) - D_0(P_w) - \Lambda(P_w)$, the time left on day w, for choices 1 or 2, resp. monotonically decreasing in $D_0(P_w) + \Lambda(P_w)$, the time spent on day w, for choice 3.

Our algorithm thus has the following structure:

1. Compute $D_0^*(\hat{P})$.
2. Assign the demand points with $\hat{f}_p = 3$ or $\hat{f}_p = 5$.
   2a. If points with $\hat{f}_p = 3$ exist: add points with $\hat{f}_p = 2$ and if necessary with $\hat{f}_p = 1$ to the days 2 and 4 until all five days have the same number of customers. The ordering of the points is based upon $(\pi_{p2} - \pi_{p1})$;
   2b. If no points with $\hat{f}_p = 3$ or $\hat{f}_p = 5$ exist: assign seed customers to days of the week using a special procedure the details of which are omitted.
3. Assign the remaining demand points according to decreasing frequency class. Within each frequency class, the next destination to be assigned is the one for which the best (weighted) sum of proximity values over all permitted day combinations is maximal, resp. minimal, among all unscheduled points. This assignment is split up into two phases as described above.
4. For the demand points p in the highest frequency class with up to now unassigned points, we compute

$$\min\{\sum_{w \in W} (D_0(P_w+p)+\Lambda(P_w+p)) \mid \text{feasible day combinations } W\}.$$

   We assign the demand point for which the expression above is minimal to the day combination which attains the minimum.
   Step 4 is repeated until no more customers are unassigned.

APPENDIX 3
DATA COLLECTION

In this appendix we will make some remarks on the preparatory work needed to undertake a case study as described in Section 6. We omit the discussion concerning the collection of the various cost parameters and the major operational accounting effort to establish the production and throughput cost functions.

With respect to the customer data, a record has to be created for each customer containing his characteristics, i.e. a specification of his address, e.g. via street and place of residence, his annual turnover in each of the commodities, and his annual number of deliveries.

Next a list of potential locations for DC's and a list of nested depot sets has to be generated, including all cost parameters for the locations concerned.

In order to generate distances and travel times between each pair of locations (of customers, depots and DC's) the following operations have been executed:

1. The alphanumerical location specifications are transformed into *coordinates* on a map. First a critical distance d_{max} is chosen to distinguish between *cities* and *towns*, the former having a diameter exceeding d_{max} and the latter having a diameter at most equal to d_{max}. A town is represented as a single point on the map, using the coordinates of its center; in the Netherlands these can be obtained from registers in special atlases. For a city representation by a single point is considered to be too inaccurate and street maps are used to obtain more precise coordinates for the delivery points within its boundaries.
2. Especially in the Netherlands, an urbanized and densely populated country, crossed by natural barriers as rivers, lakes and sea areas, Euclidean distances are far too inaccurate for most purposes. As a consequence we chose to use a *road network* constructed in order of the Ministry of Transportation [RIJKSWATERSTAAT 1974] to compute the distances between the various locations. In this network a road is represented as a series of links, each link having a length, a level number and a speed code; the level number describes the type of road (from expressway to country road) to which the link belongs; the speed code characterizes the average vehicle velocity on the link. In our study we selected the levels 1 up to 5 of the road network (out of a total of 15 levels) yielding a directed network with 5000 nodes and 11000 directed links. The low average degree of the nodes is due to the fact that most nodes merely serve to indicate speed code mutations, e.g. when a four lane highway changes into a three lane highway. Our particular case study required additional adaptations of the thus created network, e.g. the elimination of links that are forbidden for the transportation of explosive material. Finally for each

speed code the average velocity of the company's trucks on links with that speed code was computed and attached to that code.

3. Before being able to compute shortest distances between locations we need to attach each location, specified by a pair of coordinates, to one or more nodes in the network. Experience with similar studies has pointed out that a single *feeder link* for each location leads to inaccurate distance estimates. We therefore connected each location to the closest node in each of the four quadrants in the plane with this location as its origin. Next we eliminated the feeder links that cross natural barriers like lakes and rivers, a procedure that had to be done manually.
4. Finally the computation of 1.000.000 shortest distances and travel times, or cost weighted combinations of the two, between all origin-destination pairs in a set of 1000 locations is very time consuming, especially in a network of this size. However substantial computational savings can be obtained from the observation that the distance between remote customer or depot locations is never needed. Hence it is possible and advisable to confine oneself to origin-destination pairs with distances below some critical value, at least for locations of customers and depots. In a reaching type of algorithm like Dijkstra's shortest path algorithm [DIJKSTRA 1959] such an upper limit can easily be implemented, even dependent on the origin location, by terminating the generation of the shortest path tree as soon as the distances to the remaining nodes all exceed this limit. This simple trick reduced the computation time from an estimated time of 1 hour to generate the entire distance matrix to 23 minutes. Moreover to exploit the sparsity of the network we implemented a special version of Dijkstra's algorithm using heaps of bounded height [JOHNSON 1977].

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to the executive management of AGA Gas B.V., in particular to its president Dr.Ir. R. Bergsma and to Messrs. A.F.J. Buchwaldt and R. Hoorn who guided us in the day to day development and implementation of the model. In addition we are indebted to Gerard Wanrooy and Leander Jansen for their participation in large parts of the project.

REFERENCES

B.T. BAYLISS (1965) *European Transport*. Mason, London.

E.J. BELTRAMI, L.D. BODIN (1974) Networks and vehicle routing for municipal waste collection. *Networks* 4, 65-94.

J.F. BENDERS (1962) Partitioning procedures for solving mixed-variables programming problems. *Numer. Math.* 4, 238-252.

R.B.S. BROOKS, A.M. GEOFFRION (1966) Finding Everett's Lagrange multipliers by linear programming. *Operations Res.* 14, 1149-1153.

N. CHRISTOFIDES (1976) The vehicle routing problem. *Rev. Française Automat. Informat. Recherche Opérationelle* 10. 2, 55-70.

N. CHRISTOFIDES (1977) A short guide for VERSA. Mimeo.

N. CHRISTOFIDES, S. EILON (1969) Expected distances in distribution problems. *Operational Res. Quart.* 20, 437-443.

G. CLARKE, J.W. WRIGHT (1964) Scheduling of vehicles from a central depot to a number of delivery points. *Operations Res.* 12, 568-581.

G. CORNUEJOLS, M.L. FISHER, G.L. NEMHAUSER (1977) Location of bank accounts to optimize float: an analytic study of exact and approximate algorithms. *Management Sci.* 23, 789-810.

E.W. DIJKSTRA (1959) A note on two problems in connection with graphs. *Numer. Math.* 1, 269-271.

S. EILON, C.D.T. WATSON-GANDY, N. CHRISTOFIDES (1971) *Distribution Management: Mathematical Modelling and Practical Analysis*. Griffin, London.

L.B. ELLWEIN (1970) Fixed charge location-allocation problems with capacity and configuration constraints. Technical Report 70-2, Department of Industrial Engineering, Stanford University.

M.L. FISHER, R. JAIKUMAR (1978) A decomposition algorithm for large-scale vehicle routing. Report 78-11-05, Department of Decision Sciences, University of Pennsylvania, Philadelphia.

R.L. FRANCIS, J.M. GOLDSTEIN (1974) Location theory: a selective bibliography. *Operations Res.* 22, 400-410.

M.R. GAREY, D.S. JOHNSON (1979) *Computers and Intractibility: A Guide to the Theory of NP-Completeness*. Freeman, San Francisco.

A.M. GEOFFRION (1974) Lagrangian relaxation for integer programming. *Math. Programming Stud.* 2, 82-114.

A.M. GEOFFRION (1976) Better distribution planning with computer models. *Harvard Business Rev.* July-August 1976, 92-99.

A.M. GEOFFRION, G.W. GRAVES (1971) Multicommodity distribution system design by Benders decomposition. *Management Sci.* 20, 822-844.

A.M. GEOFFRION, G.W. GRAVES, S. LEE (1978) Strategic distribution system planning: a status report. In: A. HAX (ed.) (1978) *Studies in Operations Management*. North-Holland, Amsterdam, 179-204.

B.L. GOLDEN, T.L. MAGNANTI (1977) Deterministic network optimization: a bibliography. *Networks* 7, 149-183.

B.L. GOLDEN, T.L. MAGNANTI, H.Q. NGUYEN (1977) Implementing vehicle routing algorithms. *Networks* 7, 113-148.

D.P. HERRON (1979) Managing physical distribution for profit. *Harvard Business Rev.* May-June 1979, 121-132.

S.K. JACOBSEN, O.B.G. MADSEN (1978) On the location of transfer points in a two-level newspaper delivery system - a case study. Research Report 5/1978, IMSOR, The Technical University of Denmark, Lyngby.

D.B. JOHNSON (1977) Efficient algorithms for shortest paths in sparse networks. *J. Assoc. Comput. Mach.* 24, 1-13.

D. KLINGMAN, R.RUSSELL (1975) Solving constrained transportation problems. *Operations Res.* 23, 91-106.

A.A. KUEHN, M.J. HAMBURGER (1963) A heuristic program for locating warehouses. *Management Sci.* 9, 643-666.

J.K. LENSTRA (1977) *Sequencing by Enumerative Methods*. Mathematical Centre Tracts 69, Mathematisch Centrum, Amsterdam, 63-76.

S. LIN (1965) Computer solutions of the traveling salesman problem. *Bell System Tech. J.* 44, 2245-2269.

S. LIN, B.W. KERNIGHAN (1973) An effective heuristic algorithm for the traveling salesman problem. *Operations Res.* 21, 498-516.

D.H. MARKS, J.C. LIEBMAN, M. BELLMORE (1970) Optimal location of intermediate facilities in a transshipment network. Paper R-TP 3.5 presented at 37th National ORSA Meeting, Washington, D.C..

T.G. MAIRS, G.W. WAKEFIELD, E.L. JOHNSON, K. SPIELBERG (1978) On a production allocation and distribution problem. *Management Sci.* 24, 1622-1630.

MINISTRY OF TRANSPORT (1959) The transport of goods by road, Report of sample survey made in 1958. H.M.S.O., London.

MINISTRY OF TRANSPORT (1964) Survey of road goods transport 1962. Final results part I. H.M.S.O., London.

MINISTRY OF TRANSPORT (1965) Carriers' Licensing. Report of the Committee. H.M.S.O., London.

RIJKSWATERSTAAT (1974) Basisnetwork-programmatuur-reistijden (Basic network-programs - traveling times; in Dutch). Rijkswaterstaat Dienst Verkeerskunde, Centrum voor Vervoersplannen, The Hague.

R. RUSSELL, W. IGO (1979) An assignment routing problem. *Networks* 9, 1-17.

R.T. SAMPSON, M.T. FARIS (1966) *Domestic Transportation*. Houghton Mifflin, Boston, 51.

D.M. STEIN (1978) An asymptotic, probabilistic analysis of a routing problem. *Math. Oper. Res.* 3, 89-101.

TFD (1977) Distribution planning using mathematical methods. (in Swedish). Report 1977, 1, Transportforskningsdelegationen, Institutionen for Tillempad Matematik, Stockholm.

UNITED NATIONS (1967) Annual Bulletin of Transport Statistics for Europe, 1966.

C.D.T. WATSON-GANDY, P.J. DOHRN (1973) Depot location with van salesmen - a practical approach. *Omega* 1, 321-329.

M.H.J. WEBB (1968) Cost functions in the location of depots for multiple-delivery journeys. *Operational Res. Quart.* 19, 311-320.