OPERATIONS RESEARCH CENTER

working paper

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

The Mathematical Analysis of Multilocation Distribution Systems: A Survey

Uday S. Karmarkar

OR 025-73 September 1973

Supported in part by the Office of Naval Research under contract 67-A-0204-0076

MATHEMATICAL ANALYSIS OF DISTRIBUTION SYSTEMS: (A Survey)

Introduction:

This paper is concerned with multilocation inventory theory -- the study of distribution systems involving more than one inventory location. Such systems are often called "multiechelon" in the literature, but we shall here use that term in a more restricted sense. An excellent and extensive survey of work in this area was given by Clark in [30] and this paper of necessity leans heavily on that account. The main difference here is that while the survey by Clark was organized on the basis of the methodology applied, we will emphasise the nature of the problem being solved and the relation between the type of problem and the methodology used.

CHARACTERISTICS OF MULTILOCATION SYSTEMS :

In order to clarify usage of terminology and to provide a framework in which to examine applications, we discuss below some of the major characteristics of distribution system models. We will simultaneously develop some notation that will enable detailed formulation of the problem in a network flow framework. While it is not intended to create a definitive taxonomy, a modest attempt has been made to systematise and reconcile existing usage of terms.

1). Basic elements of a Distribution System :

Locations : A location is any physical point at which inventories are held. The terms "activity", " facility" and "station" are also commonly used in the literature to denote an inventory location.
We will use the term demand location to signify a location which experiences demands arising exogenously. Similarly a source location is one that is supplied exogenously. A multilocation system by definition consists of more than one location which we will index by i= 1,2,...,N.

Products : The word "commodity" is often used interchangeably with product. In general realistic problems will involve inventories of more than one product with interactions usually arising due to competition for limited resources such as inventory capacity. In the literature however, most of the work done has been in the single product case with the papers of Ignall and Veinott [17] and Patel & Karmarkar being two exceptions. We will index products by j= 1,2,..,M.

Time periods : Most of the published work in multilocation systems has been ina periodic review setting where decisions are made at fixed points in time. We shall index time periods by t= 1,2,...,T where T is the number of time periods considered in the finite horizon case.

Sources : Several exogenous sources may supply the distribution network through the source locations of the system. Exogenous sources will be indexed by s= 1,2,...,S.

Once the locations, time periods and products of a system have been labelled, we can specify particular combinations using the indices. Thus the ordered pair (i,j) for example, denotes location i inventory of product j.

The generic variable x will be used to denote flows of material endogenous to the system so that

$_j x_{i_1 t_1}^{i_2 t_2}$: Amount of product j transferred from location i_1 in period t_1 to location i_2 in period t_2.

The variable z will denote exogenously supplied material

$_j z_{s t_1}^{i t_2}$: Amount of product j supplied from source s in period t_1 to location i in period t_2.

Clearly depending on the structure of the particular problem being studied it may or may not be possible for all such variables to be non-zero. We will furthermore use the variable d to denote exogenous demands and the variable y to denote stock levels at locations so that :

$_j d_{it}$: Demand for product j experienced at location i in time period t.

$_j y_{it}$: Total stock of product j at location i at the start of period t after receipt of all shipments.

2) Structure of Distribution Systems :

Given the elements comprising the distribution system, the structure of the system refers to the type of shipments it is possible to make from a purely physical point of view. In other words in network terms, if locations are the nodes of the network, the structure of the system is to be specified in terms of the arcs of the network. Furthermore ir specifying the physical structure of a system, we ignore the time dimension and consider a time-slice or snapshot of the system in one time period to concentrate on the feasibility of shipments between actual locations . The major categories of system structure are described below :

a) General Network :

If shipment can be made between any two locations in either direction the model will be called a general network model

b) Multistage/Acyclic Network :

A network is acyclic if it has no cycles, i.e. it is not possible to make a sequence of shipments such that material starts from and ends at the same location. In the context of distribution systems we shall refer to such networks as multistage networks. Each location in such a system is assigned a stage number and a stage consists of all locations having the same stage number.

The assignment of stage numbers is done as follows:

i) Every demand location which does not supply another location is numbered 0.

ii) Every location supplying at least one i-stage location is numbered i+1.

iii) Locations supplying several stages are assigned the highest stage number consistent with ii) .

The concept of a stage stock may or may not be of practical use. There appear to be several possible alternative definitions. For example :

- the ith stage stock is the total stock held(or on order)to all locations numbered i.
- the ith stage stock is the total stock held at (or on order to) all locations with stage numbers equal to or smaller than i.

Of these two the latter is probably preferable. We can also define the stage stock at a particular location as

- the stage stock at any location is the stock held at (or on order and in transit to) that location and all other locations supplied by it. This idea will tally with our later notion of an"echelon stock".

c) Multisource network
This term refers to systems supplied by more than one exogenous source. The sources may be"plants" or "vendors".

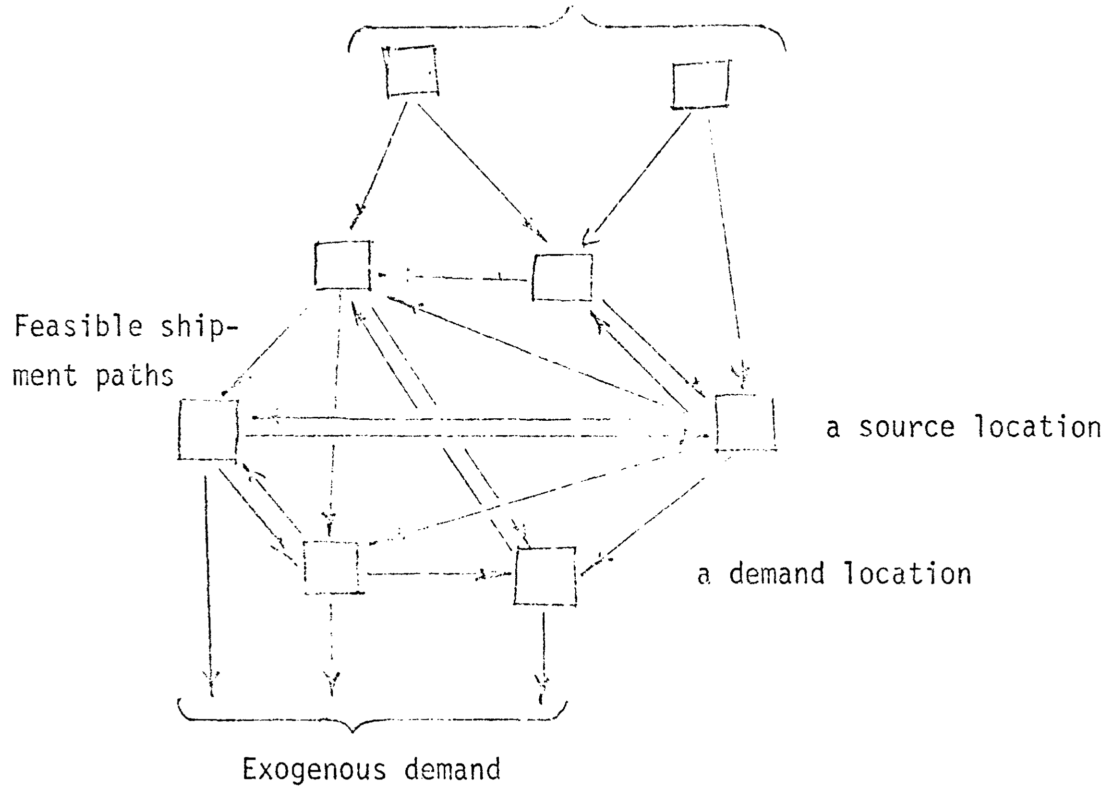
d) Delta Network :
This is a system in which 0-stage locations are the only demand locations.

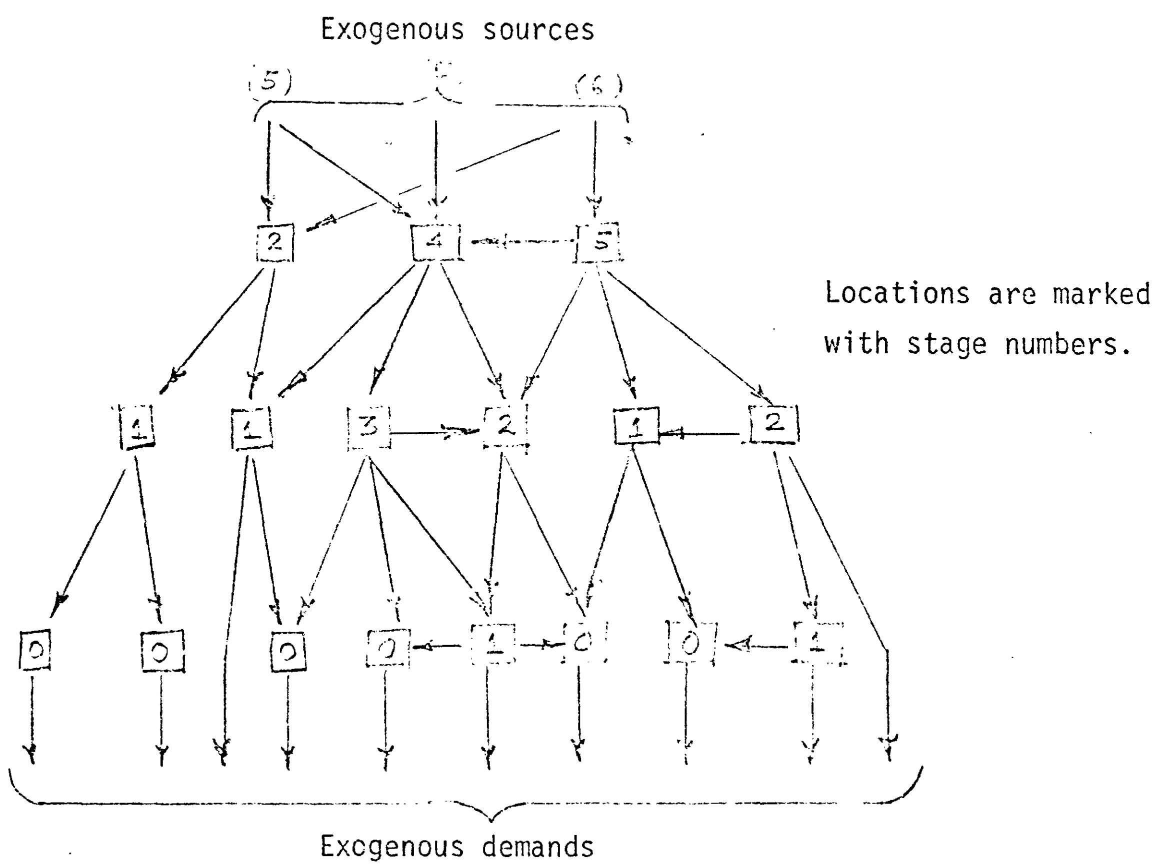
e) Multiechelon/Arborescent Networks :
An arborescence is a network in which every location is supplied by only one other location. We note that arborescences are acyclic single-source networks. Again in the context of distribution systems , multistage arborescent networks will be called multiechelon systems .
Each location in a multiechelon system is assigned an echelon number which is the same as the stage number for that location. Echelon stock at any location is the stock held at (on order or in transit to) that location and all other locations supplied by it. Total ith echelon stock is the total echelon stock at all ith echelon locations. The concept of echelon stock was proposed by Clark [ ] and was shown to be of significant use by Clark & Scarf in [5].
We distinguish two special cases of multiechelon networks that have been noted by Bessler and Veinott [3]. :

i) Wheel Network : This is a single echelon multilocation system
ii) Series Network : This has a single location in each echelon. (only one location with a given echelon number).

As we have stated earlier, the structure of the distribution system refers to the physical structure and ignores the time dimension. The physical network extended in time by replication also constitutes a network to which terms such as acyclic and arborescent apply equally well. However the terms multistage and multiechelon will be reserved to describe the physical structure of the system. It is noted in passing that an arborescent physical structure does not imply that the time extended network is arborescent.

GENERAL NETWORK

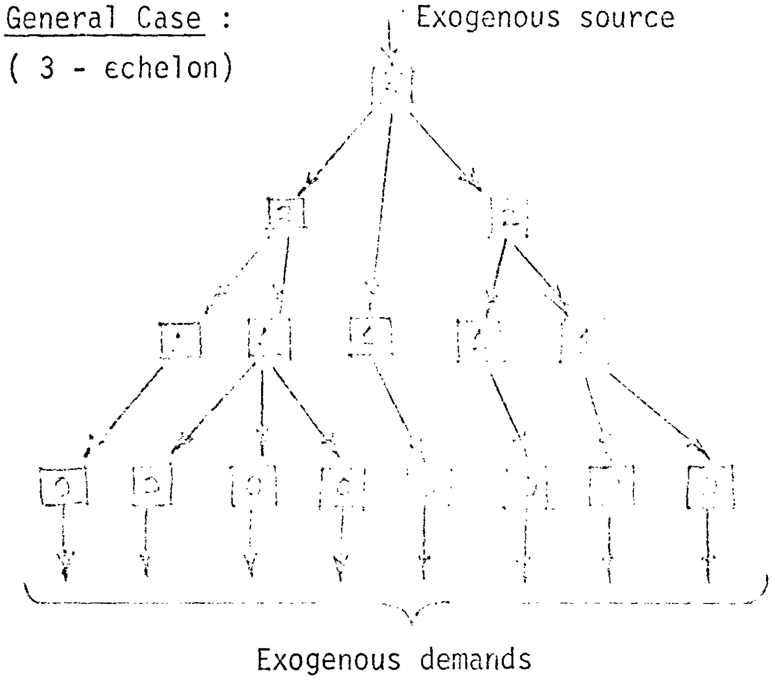




MULTISTAGE (ACYCLIC) NETWORK : 5 - STAGE.

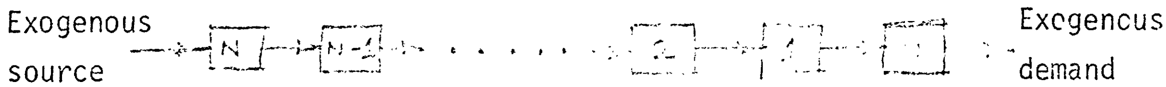
MULTIECHELON (ARBORESCENT) NETWORKS

a) General Case :
(3 - echelon)

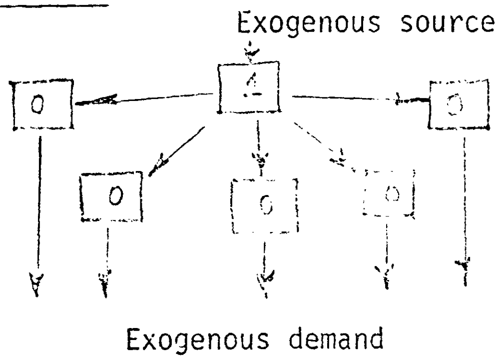


Locations are marked with their echelon numbers.

b) Series Network :
(N-echelon)



c) Wheel Network :



3) Temporal Structure of Networks :

Unlike the physical structure of distribution systems which has received considerable attention in the literature, temporal relationships have not been extensively explored. It has already been noted that arborescence is not expected to be a general property of the extended network. However if the physical structure is acyclic then under mild restrictions the extended network will also be acyclic. For example it is required that exogenous demand be satisfied before endogenous shipments are made out of a location. Acyclic time extended networks have been extensively studied by Zangwill [20] in a deterministic demand context and we will use some of his ideas in our formulation.

The temporal factor that has received most attention in multilocation theory is shipment lags. As in the classical single location case, the state space that has to be considered in say a dynamic programming approch, is greatly enlarged if lags exist. From our network point of view, the state of a system is specified by the set of arcs that connects the physical network in one period to its replicate in the next period. We shall call these interperiod arcs .

The complete specification of the structure of the distribution system thus requires extending the set of locations(nodes) in time by replicating the set for each time period, and then specifying all the possible flows of material (arcs) that can take place. In particular, interperiod arcs consist of :

$_jX_{i,t}^{i,t+1}$: Inventory of product j held at location i in the t'th period.

$_jX_{i,t+1}^{i,t}$: Demand for product j backlogged at location i in the t'th period

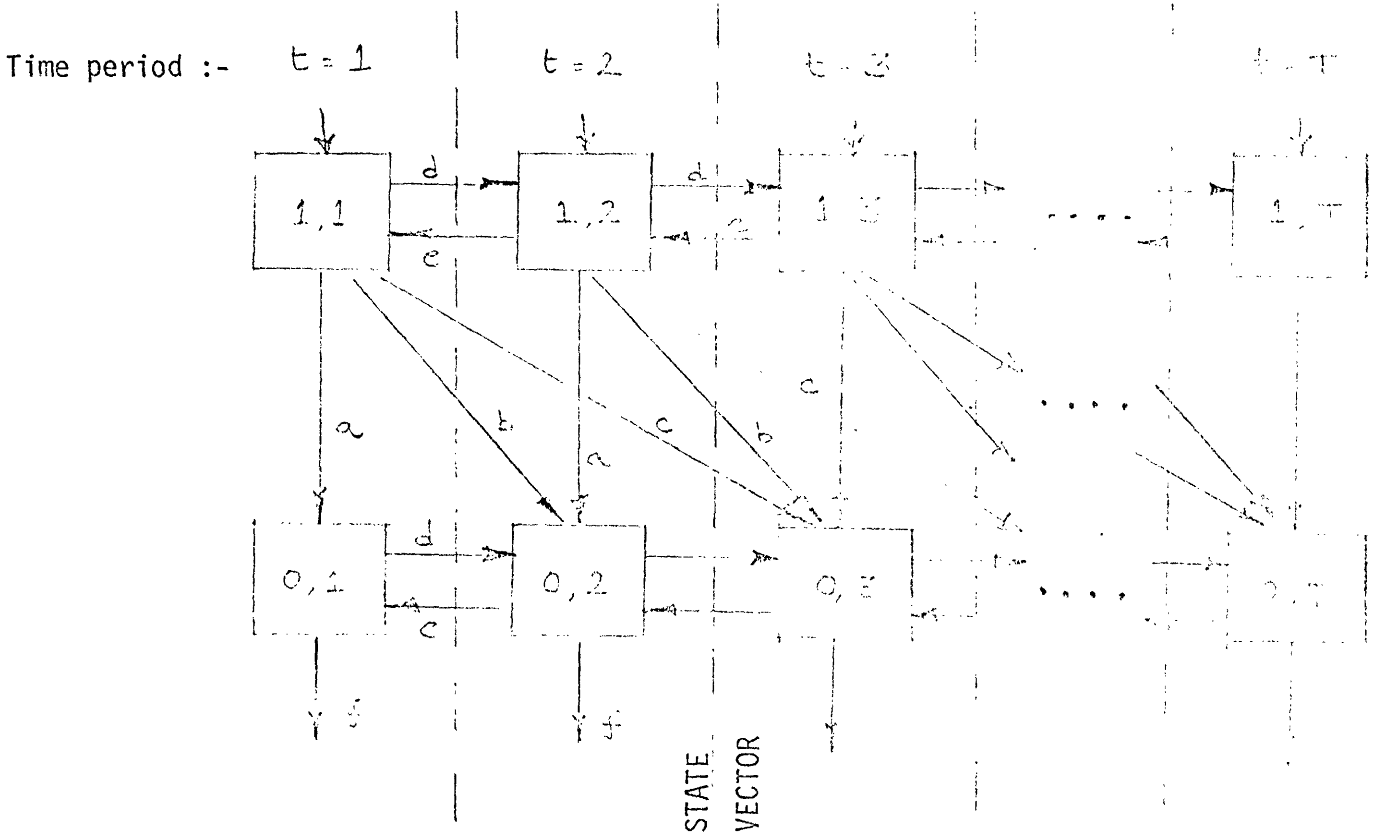
$_jX_{i_1,t}^{i_2,t+\lambda}$: Shipment of product j ordered in period t from location i_1 to i_2 with a lag of λ time periods.

In special cases it may be desired to permit demand at a location to be satisfied by a shipment from another location with some time lag. This could be incorporated as a flow

$_jX_{i_2,t_2}^{i_1,t_1}$: Demand for product j in location i_1 in period t_1 satisfied by a shipment from location i_2 in period t_2.

AN EXAMPLE OF A TIME-EXTENDED NETWORK :

Two locations in series; Single echelon



Network paths :

a) Stock distribution in current period
b) Shipment with one period lag.
c) Shipment with two period lag.
d) Inventory holding.
e) Demand Backordered
f) Exogenous demand.

4) Policy :

The term policy is used to describe the permissible alternative courses of action available in making a decision. This issue is clearly closely related to the structure of the system in terms of both the locations (nodes) as well as the material flows(arcs). The main policy areas are:

Stocking Policy :

Given initial inventories in any period, the stocking policy specifies the desired stock level at locations in terms of target stock levels for each location. The target level if feasible can be reached by

i) Ordering (exogenously) and/or by

ii) Redistribution of stock or endogenous shipments.

Clearly the permissible options determine the arcs of the network model.

Supply Policy :

When an exogenous demand is realized at a location, it may - depending upon the case in question - be partially or fully realized in several ways. Veinott in [[illegible]] has formalised this notion in terms of the function $s(y_t, d_t)$ that represents supply policy in the tth period in terms of stock remaining at a location, given stock on hand y_t and a realised demand d_t. In terms of our network framework supply policy can be represented in terms of permissible flows, node balances and(if required) by the creation of dummy nodes. The usual possibilities are :

i) Satisfy demand upto the level of stock on hand only. This is the lost sales case and requires a dummy supply source at each location where the cost of supply is set at the cost incurred in lost sales.

ii) Backlog excess demand to the next period. This is represented in a network as an interperiod flow from the succeeding period to the current period (backward in time).

iii) Expedite orders if backlogging occurs. In cases where "normal" shipments have long lead times, the possibility of speeding up shipments at the expense of added costs may exist. This simply involves the addition of the appropriate arcs to the network.

iv) Satisfy demand by direct shipment from another location. This case has been discussed above.

v) Recovery and repair of items. A significant class of problems of interest is the case where demand occurs due to item failure, and where partial or total salvage of the item involved. Depending on the time scale involved, this situation could also be thought of in terms of stocking policy. For example if recovery of items takes one time period, then the stock in the succeeding time period is simply augmented by the fraction of demand in the current period that is recoverable. We note however that in this case the usual flow interpretation of the network does not hold. (This situation c an be gotten around if we create dummy supply sources which supply probabilisticly specified amounts which must be accepted.)

Stock Review Policy :

The two usual possibilities are :

i) Periodic Review

ii) Continuous Review.

As we have stated earlier, most of the work done in this area assumes a periodic review policy with the work of Hadley & Whitin [ ] being a major exception. The notation developed so far assumes periodic review policies and the assumption is thought to be quite realistic , especially in the case of centralised control.

5) Capacity Restrictions :

There are three important types of capacity restrictions;

Supply: $_j\bar{Z}_{s,t}$: The maximum amount of product j available from source s in time period t.

Arc : $\bar{X}_{i_1,t_1}^{i_2,t_2}$: The maximum allowable flow from location i_1 in period t_1 to location i_2 in period t_2 (in terms of some common unit of capacity)

Node: $\bar{Y}_{i,t}$: The maximum holding capacity of location i in period t (in terms of some common unit of capacity).

The latter two capacities are especially significant in multiproduct situations since they are the source of interactions between products.

6) Cost Structure :

The costs which are usually considered in inventory models can be associated very conveniently with arcs of the network. The major costs involved are:

i) Ordering Cost :The cost of ordering from external sources usually consists of a purchase cost plus a shipping cost. This cost may vary with the time period in which the order is placed, the expediency of the shipment as well as the source of supply. The cost may also vary stochastically.

${}_jC^{\tau}_{s,i}({}_jx^{i,t+\tau}_{s,t})$: Cost of shipment of product j from source s in period t_1 to source location i with a procurement lag of τ time periods.

ii) Transshipment cost : This cost is incurred when there is shipment of material within the system as in the redistribution of stock or in the distribution of stock to lower stages or echelons of the system. In general the cost depends on the source, the destination, the amount shipped and the time of shipment.

${}_jC^{i_2,t_2}_{i_1,t_1}({}_jx^{i_2,t_2}_{i_1,t_1})$: Cost of shipping product j from location i_1 in period t_1 to location i_2 in period t_2.

iii) Inventory Holding costs : This is the cost incurred in holding stocks of inventory at a location. The cost is assumed to be proportional to the number of periods for which inventory is held, so that we can always write inventory costs in terms of a one period holding cost.

${}_jC_{i,t}({}_jx^{i,t+1}_{i,t})$: One period cost of holding inventory of product j at location i in period t.

iv) Backorder or Shortage Costs : Depending on the policy used, a cost is charged to demand backordered or for a shortage of stock leading to unsupplied demand. The cost is thus a function of stock on hand as well as demand realized.

${}_jC_{i,t}({}_jy_{i,t}, {}_jd_{i,t})$: Cost incurred in period t at location i due to shortage or backordering of demand for product j.

These costs may be linear, involve a fixed charge, or in general be concave functions of their arguments. The nature of the costs has important implications for the methodology employed in solving the problem.

7) Stochastic Factors :

The major source of stochasticity in distribution system models is uncertainty about exogenous demand. Depending on the circumstances, the state of knowledge about demand may be assumed to be any of the following:

i) Demand in each period is known (deterministic). The stochastic case is sometimes reduced to this case by using a safety stock to allow for the uncertainty. In this paper we are not concerned with this approach.

ii) Demand has the same distribution in each period (stationary case).

iii) Demand has a specified distribution in each period , which in general varies from period to period (non-stationary case).

iv) Demand distributions are not independent. The dependencies may be across time, between products or amongst locations.

v) The distribution of demand in each period is unknown so that it is estimated or a prior distribution is assigned and updated as information becomes available.

Other stochastic factors may be present. Significant among these are stochastic prices and cost parameters and uncertainties in shipment and procurement lag times. We will not concern ourselves with these factors since they are not treated in the current multiechelon system literature.

FORMULATION OF THE MULTILOCATION PROBLEM :

Armed with the notation we have developed, we can despite its unwieldy nature,proceed to use this notation to formulate a fairly general version of the Multilocation distribution problem. This formulation requires the specification of an objective function and a set of constraints.

The constraints can be stated immediately as :

i) Material balance at locations :

$$\sum_{i_1 \neq i} \sum_{t_1=1}^{t} {}_j x_{i_1 t_1}^{i\,t} - \sum_{i_2} \sum_{t_2}^{N} {}_j x_{i,t}^{i_2 t_2} + {}_j x_{i,t-1}^{i,t} - {}_j x_{i,t}^{i,t-1}$$

Shipments in — shipments out — Inventory in — previous backorder

$$+ {}_j x_{i,t+1}^{i,t} - {}_j x_{i,t}^{i,t+1} + \sum_{s} \sum_{t_1}^{t} {}_j z_{s,t_1}^{i,t} - {}_j d_{i,t} = 0$$

current backorder (or lost sales) — Inventory held — Exogenous supply — demand

$$\forall i, j, t$$

Node Capacities :

$${}_j Y_{i,t} = \sum_{i_1 \neq i} \sum_{t_1=1}^{t} {}_j x_{i_1 t_1}^{i\,t} - \sum_{i_2} \sum_{t_2}^{N} {}_j x_{i,t}^{i_2 t_2} + {}_j x_{i,t-1}^{i,t} - {}_j x_{i,t}^{i,t-1} \leq {}_j \bar{Y}_{i,t}$$

Stock held after shipments — Shipments in — Shipments out — Inventory in — Backorder

$$\forall i, j, t$$

Total inventory
$$\sum_j a_j ({}_j Y_{i,t}) \leq \bar{Y}_{i,t}, \quad \forall i, t$$

Arc Capacities;

$$\sum_j a_j ({}_j x_{i_1 t_1}^{i_2 t_2}) \leq \bar{X}_{i_1 t_1}^{i_2 t_2} \qquad \forall i_1, i_2, t_1, t_2$$

where a_j's are coefficients of conversion to some common unit of capacity.

The objective function is simply to minimize the total expected value of the costs associated with the flow in each arc. It should be noted that in order to be able to take the expectation, we must prespecify a supply policy, that fixes the decision variables in each period, given the starting conditions in that period. The detailed formulation of this concept can be found in the papers by Veinott [20] and Bessler and Veinott [3].

THE NATURE OF THE PROBLEM :

The problem as formulated has some important characteristics that bear upon the methodology applied as well as upon the characterisation of the solutions obtained. The problem has linear constraints so that it has the features of a network flow or at least a linear programming problem. At the same time the costs involved are expected costs of the newsboy type, sothat the problem has the character of a newsboy or (s,S) type of situation. Furthermore the multiperiod stochastic form of the problem suggests that the problem can be thought of in terms of stochastic processes.

An intuitive way to visualise the problem in terms of the decision making process is as follows :(Gross [10], Patel & Karmarkar [21]).

i) In each period, stocking decisions are made with respect to redistribution of stocks and exogenous ordering. This is in general a multicommodity capacitated flow network situation and the costs associated with these decisions are deterministic. The result of the decision is to achieve some target stock levels at each location and the cost of reaching the target levels depends on the starting stocks before the decision.

ii) Interperiod arcs for a given location represent realization of exogenous demand and inventory held, shortage, or demand back-ordered with their associated costs. These costs depend on the target stock levels achieved and the demand realized. Since demand is stochastic these costs are stochastic and are of the newsboy type. (Other interperiod arcs across different locations represent lagged shipments with deterministic costs).

iii) At the start of the next period a stocking decision is to be made again. It is clear that the starting stock is a random variable which depends on the stocking decision in the previous period and the demand distribution in the previous period. Thus the problem has the character of a stochastic process.

Thus the problem is essentially a continuous state, discrete time parameter Markovian decision problem where the stocking decision corresponds to the choice of a transition matrix, the state of the system is the vector representing stock levels at each locations, and the realization of demand constitutes a transition. If shipment lags exist, the dimensionality of the state space is augmented to include all interperiod arcs.

The most general version of the multilocation problem has not been solved in the literature. Some simplifications are usually made and the method of attack usually reflects the simplifications . The following section briefly reviews the important papers published in this area.

MAJOR APPROACHES TO MULTILOCATION PROBLEMS :

I). General Network : Single product, Single period,uncapacitated problem, with linear costs.

This problem was formulated by GROSS in [9].The method of attack involved a detailed examination of optimal policy under different starting stock conditions, and a complete solution was obtained for the two location case. Despite the seemingly sweeping simplifying assumptions, this was the first attack on the general network problem and demonstrated that the complete solution to the problem was quite complex. In fact the procedure used proved to be too complex for the n-location case and Gross suggested that search procedures be used which would involve in general,n^2 ordering and shipping decision variables.

KRISHNAN & RAO in [19] have tackled a one period problem similar to that proposed by Gross. However, while Gross' approach considered ordering and shipping decisions made simultaneously at the start of the period, the approach here was to determine optimal ordering decisions given that transshipment decisions could be deferred till demand was realized.An additional simplification made in this paper was to assume that all transshipment costs are equal. This allows arbitrary partitioning of the locations into groups with the same transportation cost obtaining between any two groups. This assumption is critical to the solution method which iteratively partitions the locations into 1,2,...,n groups successively. Furthermore an assumption of Normality of demand distributions was made which greatly simplifies the computation of the optimal policy due to the additivity properties of the Normal distribution.

PATEL & KARMARKAR in [ ] have also studied the one period general network problem as formulated by Gross. Their approach was to decompose the problem into a stocking decision that is a transportation problem , and decoupled newsboy problems that represent the realization of demand with the associated holding and shortage costs. This approach leads to a characterization of optimal policies in terms of the dual of the transportation subproblem. Specifically it is shown that there is a correspondence between the optimal policies and the extreme points , edges , faces etc. of the dual . This method is not suitable for the numerical solution of large problems, but the exact solution provided by Gross for the two location case is easily recovered, and it is shown that the three location case involves 37 policies, as compared to seven for the two location case. For the numerical solution of large problems, the problem has been formulated as an LP with column generation.This approach is quite robust in the sense that it is easily extended to incorporate capacity constraints and the multiproduct case. The major deficiencies in this approach are thus the linear cost and the single period assumptions.

II) Multiechelon (Arborescent) Systems : Single product, Multiperiod

CLARK & SCARF in [5] and [6] have examined multiechelon systems with the intention of establishing the possibility of decomposing the n-variable problem into n single variable problems for which the usual (s,S) policies obtain. The procedure uses the concept of an echelon stock which is the total stock held at,on order or in transit to a location and all other locations supplied by it. Costs are then represented in terms of echelon stocks at each echelon with a penalty term added to account for inability to supply lower echelons. This procedure is proved to be optimal in the special case of a series network where the costs at the lowest echelon are of a fixed charge type and at higher echelons of a linear type. Approximate solutions are obtained for the general fixed charge case. The procedure runs into difficulties in the general multiechelon case in specifying the costs of not supplying lower echelons, since these costs depend not only on the shortage of stock but also on the way the available stock is apportioned.

Another proof of the optimality of the Clark-Scarf procedure for the series network case was given by Veinott in [ ] using a theorem of concave programming by Karush [ ]. A similar proof was given by IGLEHART & MOREY in [ ] in considering the case of two location series situation where the Clark-Scarf approach is extended to consider the accuracy of demand forecasts.

HOCHSTAEDTER [ ] has considered a situation where two warehouses are supplied by a central warehouse with no shipment between the two warehouses (wheel network). The satellite warehouses are assumed to follow (s,S) policies that are optimal for them. An approximation for the costs of such a system are obtained in terms of upper and lower bounds on the total costs.

FUKUDA in [ ] extends the Clark-Scarf approach for the case of the series network to allow for the disposal of stocks in each period. (In our formulation of the general problem above we have not included this possibility butthe addition is simple). Linear ordering costs are assumed and it is assumed that a lag of one period is involved in supply from or disposal to a higher echelon. Stock can only be disposed of by moving it up through the system till it exits at the highest echelon.

Finally, IGLEHART & LALCHANDANI in [ ] also using the classic methods of dynamic programming used in the papers above, have analysed the case of a wheel network with two satellite locations, assuming linear costs and a limit on the total stock held at the two 0-echelon locations. The exact form of the optimal policy is derived in this case but it is indicated that extension to larger problems would require techniques beyond those used in this paper. We may note that this paper attacks the problem of allocation of scarce resources to competing lower echelon locations and thus bears upon the Clark-Scarf paper where this situation caused their decomposition to break down. The incorporation of these results into the Clark-Scarf approach remains to be investigated. This problem is also related to the redistribution and stock allocation type of problem formulation.

The second major approach to multiechelon systems has been that of Veinott and others. Clark in his survey [[illegible]] has termed this approach "dynamic process analysis".

In [[illegible]] Veinott studied a multiproduct single location situation in which the n-period optimality of one-period decision policies was examined. While this is not the situation of interest here, a similar approach is used in later papers on multilocation problems and this paper presents some of the basic ideas involved. In effect the multiproduct one-location situation is later shown to have a correspondence with the single-product multilocation situation through the notion of "subsitutability" of products.

BESSLER & VEINOTT [2] have studied problems with an arborescent network structure and linear costs. Conditions for optimality of a "basestock" type of policy are obtained. In the study of arborescence structures, backlogging is allowed only at the highest echelon, with demand being passed up the network instantly, with satisfaction wherever possible. The n-period problem is decomposed into one-period problems, each involving m variables corresponding to the base stock levels at the m locations.

Later in the paper bounds for the optimal stock levels are obtained, and iterative computational methods to sharpen these bounds are suggested. The ordering of base stock levels on the basis of the spread of demand distributions is also examined, and the effect of parameter variation on stock levels is studied.

IGNALL & VEINOTT [[illegible]] examine the optimality of "myopic" or one-period policies further. A linear cost structure is assumed. Starting stocks are to be allocated to each location at the beginning of each period and shortages at locations may be satisfied by stocks at other locations provided they are replaced from exogenous sources in the next period. Under such assumptions, the myopic policy is shown first to be optimal for networks with a "nested" structure in the stationary case. The results are later extended to the non-stationary and delivery lag cases.

We have discussed above some of the major approaches to multilocation distribution problems. There is however a significant portion of the literature that we have not touched upon. We here restrict our discussion of this work to a few remarks:

III) Redistribution & Allocation problems : These formulations consider the one time (one-period) allocation or redistribution of a fixed amount of stock through a distribution system. Redistribution problems are typically cast in a general network format and have been extensively discussed by Allen in [illegible] and [illegible]. The allocation type of formulation usually assumes a multiechelon structure often specialised to the wheel network. Work in this area includes that of Brown [4, Ch. 19,20], Simpson [illegible] and the comment by Harris [illegible]. We may note that the allocation/redistribution problem is completely subsumed by the general network problem format.

IV) Low Demand Situation : Typically the low demand problem allows certain characteristic methods to be adopted which exploit the finite demand description. Hadley and Whitin [11, 12] have used special assumptions about the distribution and have have examined a continuous review policy, which has not been done in most of the literature. Love [illegible] brings to bear what is essentially a Markov decision process formulation, which allows computational results in small problems. El-Agizy [illegible] has used a stochastic linear programming approach. Simon in [illegible] has studied the stationary properties of a two echelon system.

V) Recoverable Items : As we have stated earlier, this is an important class of problems, but one which is qualitatively different from the situation of interest here. The work of Sherbrooke [illegible] in this area is well known and extends from a theoretical analysis to an operating control system.

OPERATIONAL METHODS IN MULTILOCATION SYSTEMS

It is apparent that the state-of-the-art of theoretical approches to multilocation systems does not as yet extend to the design of optimal inventory control systems in a real world sense. Indeed, except for a couple of exceptions, this has not been the intent of most of the investigators. The main thrust of the literature examined in this area appears to be towards establishing results of a general theoretical nature or to relating the problem to simpler situations where the solutions are known with a view to characterizing the nature of the solutions. However, while many of these results may not be suited to direct application due to restrictive assumptions or computational infeasibility, they generally have important qualitative implications for the design of actual control systems. Let us briefly review some of the main results :

i) The approach of Clark, Scarf and others has been directed towards reducing the many location problem to an equivalent or near equivalent set of one location problems. The approach is restricted to single product arborescent networks and is optimal only in the series case for a special cost structure. Even in these cases the computations required can be quite extensive. The results suggest that policies based on total echelon stocks at each location with penalties for insufficient location stock are likely to be fairly good.The results also suggest a "one-for-one" or (S-1, S) policy at lower echelons and a (s,S) policy at higher echelons, with demand being satisfied in all cases upto the level of available stocks. This form of policy arises as a direct result of the cost structure assumed, with linear ordering costs everywhere, except at the highest echelon where a fixed charge structure is assumed.

ii) The approach of Veinott and others suggests that single period policies will often be optimal in n-period problems and that the single period policies will involve a target stock level or "base-stock" at each location. These results are obtained assuming linear costs, and thus cannot be assumed to be good approximations for the setup charge case, simply because in the latter case, a tradeoff exists between cost of additional inventory held in a given period versus setup costs in future periods.

Clark [[illegible]] has noted that the importance of ordering setup charges may be assumed to diminish with increasing automation and centralisation, but significant concavity of cost structure may also exist for reasons of economies of scale.

iii) The work of Gross, Krishnan-Rao and Patel & Karmarkar has studied the exact form of optimal policies in a one-period linear cost situation. The results show that in most cases only certain target stock levels need be considered in the problem as stated. Methods of computing solutions for large problems are also provided but the scope of definition of the problem is too restricted to be of direct practical use.

iv) The allocation models of Simpson [[illegible]], Brown [[illegible]], Iglehart and Lalchandani [[illegible]] and others indicated the importance of stockout criteria such as weighted probability of stockout, and of costs ratios in ranking locations for purposes of apportioning stock. It has also been shown by Bessler and Veinott that the "virtual stockout probability" at a location is related to a ratio of the "effective" shortage and holding costs.

v) A number of papers such as those of Hadley and Whitin [[illegible]], Love and others have examined the low demand case assuming particular demand distributions of finite demand. These do not in general help with optimal solutions to the more general problem formulations we have described. Nor does it seem to be easy to extract qualitative results that might be useful in a broader framework.

vi) The work of Sherbrooke [[illegible]] has been claimed to be the first instance of theoretical analysis extended to the design of an operational inventory control system for the recoverable items situation. As stated earlier however, this class of problems while of great interest, differs from the problem as formulated here and hence the results do not appear to be transferable.

Let us now examine practical methods of inventory management in multilocation systems that are in current use. There are two major ways of conceptualising the operation and co-ordinated control of distribution systems:

a) Demand-Pull : Material is drawn down into the distribution system by demand originating at demand locations.

b) Stock Allocation ("supply-push") : A fixed production volume or resource is available and must be allocated to loations within the distribution system. Redistribution of stock may also be involved.

c) Stock redistribution : A fixed total volume of product is available scattered through the system, and must be redistributed to various locations.

The first two viewpoints tend to look at the system as multistage or multiechelon, rather than as a general network. The first is more oriented to a series situation or to multiechelon networks with many echelons. The latter appears to be better suited to wheel networks or systems with fewer echelons (stages). In this connection , it may prove useful to define some index of the divergence or degree of spread of multiechelon networks. We m ight use for example some index of the fanout of the network relating to the number of locations served by one location. Thus we might define the average fanout of a multiechelon network as the number of locations divided by the number of echelons. Thus for an m-location system, the average fanout would be m for a wheel network and one for a series network.

The third conceptualization tends to look at the system as a general network and generally addresses the problem as a transportation model. In a practical application, the approach would be to somehow remove the stochastic factors in the model resulting in the usual deterministic transportation model. This viewpoint does not provide any special intuition about multistage structures, and is thus not usually a basis for practical intuition-based control system.

Practical methods of operation usually involve some sort of operating policy which may have varying degrees of centraliaation. At one extreme is the completely decentralised case, where each location is managed independently and follows its own policy bases on the demand it experiences. This is clearly a sort of lower bound in effectiveness to more centralized methods of control, since co-ordination of policies is bound to reduce cost by better allocation of resources. The completely

...tem is an example of a demand-driven system.

...nly advocated for control of multilocation ...espond to the first two conceptual models

... (Kimball [..], see also Magee & Boodman [..]).

...nd-pull type. Briefly, the system involves ...k holdings of the distribution network in setting ... production quantities and safety stocks are ...s, and the latter are used to account for the ... system. At lower levels demand is satisfied ...lable and information on demands throughout ...wards immediately to higher echelons as they ...tem is multiechelon, the sequence of locations ...s passed upwards and material drawn down by a ...quely specified. The distribution of stocks to ...ade on the basis of realized demand as reported ... the distribution decision is centralised, ...ituation with two locations competing for ...location decision is made locally. The allo- ...through the system remains to be made and this ...en the level of service desired at a particular ...ible delivery lag against the cost of holding ...ome lag is permissible, stocks can be held at ...general less stock will be required, since ...lobal demand variations, rather than local ...formation about stockout possibilities may be ...nout lists which are used to allocate scarce ...ts from lower echelons.

...respondence will be noted between the base stock ...rom multilocation theory. In particular, the ... by the one-for-one policies at lower echelons ...d demand is to be passed up the system as ... the highest echelon uses an (s,S) policy. ...direct result of the cost structure as was ...sumed that the base-stock method does not ... account. In the case of such a fixed charge structure, the base stock method might conceivably be modified to distribute stock to lower echelons in lot sizes which take setups into account. Since the Clark-Scarf approach emphasisedthe usefulness of the echelon stock concept, the base stock method might be revised to use echelon stocks in setting base stocks for lower echelon locations, and for determining echelon runout information in making allocations.

The approach of Veinott et.al. suggests that base stock levels be set on the basis of demand at individual locations. The one period "myopic" policy is a good approximation to the optimal base stock level in most cases. Furthermore, the optimal policy is such as to make the stockout probability equal to the ratio of effective shortage and holding costs.

2) Stock Allocation to Warehouses : (Brown [4], Harris [[illegible]],Simpson[[illegible]]) This method as the name suggests, conceptualises the problem as one of pushing out stocks from the plant to the distribution system. Stock allocation is typically based on criteria such as probabilities of stockouts, and runout times. In particular, Brown is concerned with the total remnant stock at all locations when the first location runs out, and Simpson shows that appropriately weighted stockout probabilities at all locations should be equalised. The paper by Iglehart & Lalchandani has obtained the exact form of the optimal policy for linear costs for a simple case. The allocation situation can also be regarded as a special case of the redistribution models such as those proposed by Allen and of the general network models of Gross and Patel & Karmarkar.

Of these two systems the base stock method is more commonly used than the allocation approach. There are several good reasons for this state of affairs. The base stock system requires a minimum of computation in actual operation, and is flexible in the sense that local adaptations and improvements can easily be made. For example, some features of the allocation concept can be incorporated in the base stock system by using runout information to allocate scarce stock to lower echelons. Furthermore, the base stock system enables production targets to be set on the basis of information of realised demand, when production rates are variable. Indeed the allocation method is appropriate where the production volume is fixed in each period, or where there is some fixed volume of nonconsumable resource to be reallocated (redistributed) at the beginning of each period.

Summary :

To conclude, most of the literature in this area suffers from one or more shortcomings as far as practical applications are concerned. Generally some form of special physical structure is assumed, with several papers examining multiechelon (arborescent) systems. Admittedly, this is not an excessively impractical class of problems, but on the other hand the results obtained do not always extend to the whole class.

Cost structures are generally assumed to be linear, and this is a serious restriction since the nature of the solutions in the fixed and concave cost cases is likely to be qualitatively different due to the tradeoffs between holding and setup costs, and setup versus shortage costs which become operative. Other serious shortcomings are the lack of capacity restrictions and the lack of multiproduct formulations. Many of the results also suffer from prohibitive computational requirements; and while these might be unavoidable in view of the complexity of the problem, some attempt could conceivable be made to simplify computation somewhat at the expense of optimality. It should be possible, and is highly desirable to extract simple approximations, heuristics or qualitative results from much of the literature but this has not been done in most cases.

From the point of view of theoretical investigations, the fixed charge case needs to be studied further. Qualitative characterisations of solutions to this problem could be valuable, even where exact solutions are not forthcoming. Investigations should also be extended to include the general network type of problem, which has not been tackled in a multiperiod context .

It seems likely to this author that any practical method of obtaining optimal or suboptimal solutions to large problems with any significant degree of interactions between variables, is going to involve techniques of mathematical programming. The value of this approach has been demonstrated in both the theoretical as well as the numerical computational area by the Patel & Karmarkar paper. What is required is an ability to tackle fairly large stochastic linear programming and network problems. This problem has been investigated by Connors & Zangwill [ ], El-Agizy [ ] and Szarc [ ] amongst others. However these are suited to finite discrete demand situations and furthermore the fixed charge problem has not been examined. Nevertheless this approach is thought to be extremely promising, especially because of its success in application to the discrete case of production and inventory problems. (Zangwill [ ]).

REFERENCES :

[1] Allen, S.G.; "Redistribution of Total Stock over Several User Locations", NRLQ, 5, 337-345 (1958).

[2] Allen, S.G.; "A Redistribution Model with Set-up Charge"; Mgmt.Sci., 8, 99-108 (1961).

[3] Bessler, S.A.; A.F.Veinott; "Optimal Policy for a Dynamic Multi-Echelon Inventory Model"; NRLQ, 13, 355-389 (1966).

[4] Brown, R.G.; Decision Rules for Inventory Management ; Holt, Rinehart and Winston Inc., New York (1967).

[5] Clark, A.J.; H.Scarf ; "Optimal Policies for a Multi-Echelon Inventory Problem"; Mgmt.Sci.,6,475-490 (1960).

[6] ________ ; _______; " Approximate Solutions to a Simple Multi-Echelon Inventory Problem"; Chap.5 in K.J.Arrow; S.Karlin, and H.Scarf (eds.): Studies in Applied Probability and Management Science ; Stanford U.Press, Stanford Calif., pp.88-110 (1962).

[7] Connors,M.M.; W.I.Zangwill; "Cost minimisation in Networks with Discrete Stochastic Requirements"; Ops.Res., 19, 794-821 (1971).

[8] El-Agizy, M.; "Dynamic Inventory Models and Stochastic Programming"; IBM J.of R&D,(1969).

[9] Fukuda, Y.;"Optimal Disposal Policies"; NRLQ, 8, 221-227 (1961).

[10] Gross,D.; "Centralised Inventory Control in Multilocation Supply Systems"; Chap.3 in H.Scarf, D.Gilford and M.Shelley (eds.): Multistage Inventory Models and Techniques; Stanford U.Press, Stanford, Calif.(1963).

[11] Hadley G.,T.M.Whitin; "A Model for Procurement,Allocation and Redistribution of Low Demand Items"; NRLQ, 8 , 395-414 (1961).

[12] Hadley G.,T.M.Whitin; " An Inventory Transportation Model with N Location"; Chap.5 in H.Scarf, D.Gilford and M.Shelley (eds.): Multistage Inventory Models and Techniques ; Stanford U. Press; Stanford Calif. (1963).

[13] Hochstaedter, D.;"An Approximation of the Cost Function for Multi - Echelon Inventory Model"; Mgmt.Sci., 16, 716-727 (1970).

[14] Iglehart D.L.;"Recent Results in Inventory Theory";J.Ind.Eng.; 18, 48-51,(1967)

[15] Iglehart D.L.; A.Lalchandani; "An Allocation Model"; SIAM J.Appl.Math. 15, 303-323 (1967).

[16] Iglehart D.L.;R.C.Morey; "Optimal Policies for a Multi-Echelon System with Demand Forecasts"; NRLQ, 18, 115-118 (1971).

[17] Ignall E.; A.F.Veinott; "Optimality of Myopic Inventory Policies for Several Substitute Products"; Mgmt.Sci.,15, 284-304 (1969).

[18] Karush W.; "A Theorem in Convex Programming", NRLQ, 6, 245-260(1957).

[19] Krishnan,K.S.; V.R.K.Rao;"Inventory Control in N Warehouses"; J.Ind.Eng., 16, 212-215 (1965).

[20] Love,R.F.;"A Two-Station Stochastic Inventory Model with Exact Methods of Computing Optimal Policies"; NRLQ, 14, 185-217 (1967).

[21] Patel, N.; U.S.Karmarkar; "The General N-location Distribution Problem"; Sloan School Working Paper; (1973)(to be published)

[22] Sherbrooke, C.C.; " METRIC: A Multi-Echelon Technique for Recoverable Item Control"; Ops.Res.,16, 122-141 (1968).

[23] Simon,R.M.; "Stationary Properties of a Two-Echelon Inventory Model for Low Demand Items"; Ops.Res.,19; 761-773(1971).

[24] Simpson,K.F.;"A Theory of Allocations of Stocks to Warehouses";Ops.Res., 7, 797-805 (1959).

[25] Szwarc, W.; "The Transportation Problem with Stochastic Demand";Mgmt. Sci., 11, 33-50 (1964).

[26] Veinott, A.F.;"Optimal Policy for a Multi-Product, Dynamic,Non-Stationary Inventory Problem"; Mgmt.Sci.,12, 206-222 (1965).

[27] Veinott, A.F.; "The Status of Mathematical Inventory Theory"; Mgmt.Sci.; 12, 745-777 (1966).

[28] Zacks,S.; "A Two-Echelon, Multi-Station Inventory Model for Navy Applications"; NRLQ, 17, 79-85 (1970).

[29] Zangwill, W.I.; "A Backlogging Model and a Multi-Echelon Model of a Dynamic Economic Lot Size Production System - A Network Approach"; Mgmt.Sci., 15, 506-527 (1969).

[30] Clark, A.J.; " An Informal Survey of Multi-echelon Inventory Theory"; NRLQ, , 621-650

[31] Harris, G.H.;" Allocation of Stocks to Warehouses"; Working Memorandum, GHH-8,Case#69973; A.D.Little Inc., Cambridge(1968).

[32] Kimball, G.E.; "General Principles of Inventory Control"; Internal Memorandum, A.D.Little Inc., Cambridge.

[33] Magee,J.F.; D.M.Boodman; Production Planning and Inventory Control; McGraw-Hill, New York(1967).

[34] Simpson, K.F.; "A Theory of Allocation of Stocks to Warehouses";OR, 6 (2); 863-873 (1958).