9

DISTRIBUTION SYSTEM DESIGN

Arthur M. Geoffrion
University of California, Los Angeles

James G. Morris
University of Wisconsin-Madison

Scott T. Webster
University of Wisconsin-Madison

9.1 Introduction

Since the 1950's, formal computer models have been used for answering distribution system design questions such as:

- How many stocking points should there be, and where should they be located? Should they be owned?
- Should all stocking points carry all products or specialize by product line?
- Which customers should be served by each stocking point for each product?
- Where should plants be located?
- What should be produced at each plant and how much?
- Which suppliers should be used and at what levels?
- What should the annual transportation flows be throughout the system? Should pool points be used, and if so where should they be?

Warehouse location problems, posed as mixed-integer programming (MIP) models with binary variables for fixed charges related to site choice and continuous variables for the flow of goods, began to appear in the operations research literature with such papers as [63] (which describes a heuristic for a nonlinear warehouse location model). About two decades ago the structural design of distribution systems using an optimizing approach became technically feasible. Geoffrion and Powers [458] provide perspective on algorithmic and associated evolutionary developments, stating that

> "It is now possible for companies of all sizes to find distribution system designs that are optimal for all practical purposes even, in many cases, when the scope of the design problem is extended to the complete logistics system in the broadest contemporary sense."

Yet a challenge remains. Although perils attend non-optimizing (heuristic) approaches [459], optimization has failed to become the dominant approach to distribution system design. The slow demise of non-optimizing approaches in actual practice is considered "a lesson in humility to the MS/OR profession [458]." Our intention is to establish a base for embracing powerful optimizing approaches by presenting a guide on tools conductive to fast analysis of distribution systems. Specifically, we wish to: (i) convey the power of highly simplified models for quick diagnosis and exploratory purposes, as well as for developing insights into cost relationships in distribution systems; and (ii) show how algebraic modeling language systems are convenient for quickly specifying a model and for solving aggregate models of distribution system design problems through access to general commercial solvers.

In order to provide a context for our discussion, we begin with a case study in the form of a series of memos that illustrates a systematic response to a question on the suitability of an existing system, from initial diagnosis through results of formal analysis. We then consider each tool in turn, beginning with a highly simplified model known as the general optimal market area model, then moving on to algebraic modeling language systems. Our emphasis is on tools that can be easily and inexpensively implemented in industry. Algorithms are not our focus. As a convenient reference for readers interested in more detail on methodology and applications, we conclude with a brief annotated bibliography. Attention is restricted to distribution planning models wherein location of facilities is a central question.

9.2 A Case Study

[memo 1]

TO: Director of Management Science
FROM: Vice President of Logistics
DATE: June 1

As a newly appointed Vice President with responsibility for our network of distribution centers, I would like your advice concerning the suitability of this network in the current business environment.

We've had the same 13 DC locations for several years now, although the volume of business has been growing and freight rates, especially on the out-bound side, have been increasing with alarming rapidity. In the most recent 12-month period, our costs have been as follows on a total volume of 25 million CWT:

Distribution center fixed	$ 6,150,000
Distribution center variable	18,500,000
Inbound transportation	42,850,000
Outbound transportation	29,100,000
	$96,600,000

Do you think it would be worthwhile to build a model of our distribution system in order to help us reduce costs? I know that you could lead such an effort, but I would need some assurances concerning the potential for cost savings and where those savings might come from.

——————[memo 2]——————

TO: Vice President of Logistics
FROM: Director of Management Science
DATE: June 10

Thank you for your inquiry concerning the possibility of a distribution network modeling study. It is always difficult to predict in advance what the cost savings will be for a major modeling effort, but I will do the best I can.

An evaluation of the merits of our current distribution network breaks down into three fundamental questions:

Q1. How well are we using our current distribution network? That is, how well are we loading plants, allocating production to the DC's, and assigning customers to DC's?

Q2. How good are our current 13 DC locations?

Q3. Do we have too many or too few DC's?

I list the questions in this order because it is the natural one not only for diagnosis, our present concern, but also for redesign of the network. We cannot think usefully about changing DC locations (Q2) until we understand how any given set of locations should be used (Q1), and we cannot think usefully about changing the number of DC's (Q3) until we understand how any given number of DC's should be located (Q2) and used (Q1). Each question is considered in turn.

Q1. How well utilized is the current distribution network?

The growth of demand has, as you know, been putting a real strain on our capacity expansion program. For most product lines, nationwide production capacity is only about 10% greater than nationwide demand and the location of this capacity is quite poorly matched to the location of demand. This poses a fairly easy problem of plant loading, as there is little latitude to decide how much of each product to make at each plant, but it does present a rather intricate problem of allocating production to DC's.

For a given set of customer-to-DC assignments, the DC's look like demand points to the plants. It is possible to calculate the plant loadings and plant-to-DC allocations for each product so as to *minimize* the sum of production and inbound transportation costs while honoring plant capacities and DC demands. This could be done easily using an optimization technique you have no doubt encountered before called linear programming. But what if the customer-to-DC assignments are changed? That would require resolving the loading/allocation problems and would also change the outbound transportation costs. Simultaneous cost minimization for loading, allocation, and customer assignment presents a very difficult optimization problem owing to the gigantic number of possible ways to assign customers to DC's. The only good news is that our DC's have sufficient unused equipment and space that customers probably do not have to compete significantly for limited DC capacity, but even so the problem is beyond the reach of the general purpose optimization software presently available.

The bottom line is this: if the joint plant loading/product allocation/customer assignment problem is so difficult for present day computers and generally available software, we should not expect that our company is solving it particularly well. Indeed, we have never taken a system-wide look at this problem. We have only nibbled at it in manual studies of limited scope. A thorough analysis would not only capture the previously inaccessible savings, but it would also capture the accessible savings that might otherwise escape notice for a year or two or even more. There is always a lag between the constant creation of savings opportunities by the turbulent business environment and our discovery of these opportunities in the normal course of running our business. A thorough analysis would eliminate this lag for the time being.

I would offer a conservative guess that, if we were to acquire the special purpose computer software to solve this problem properly, transportation costs would drop by about 2%. This savings factor seems modest enough and is in keeping, so I am told by consultants who specialize in this line of work, with the experience of other firms which have solved the joint plant loading/product allocation/customer assignment problem with tightly constrained supply on an integrated, system-wide basis for the first time. Thus I estimate that our cost breakdown would have been as follows had we utilized our current 13 DC's optimally:

DC fixed costs	$ 6,150,000
DC variable costs	18,450,000
Inbound transportation	42,000,000
Outbound transportation	28,500,000
	$95,100,000

Additional savings would probably accrue for production costs, but I will omit consideration of that for the sake of simplicity.

Q2. How good are the current 13 DC locations?

A small random sample of outbound freight rates shows that delivery costs average $0.0075 per CWT-mile. Total annual delivered volume is 25 million CWT, of which about $\frac{1}{4}$ is in the very cities in which our DC's are located. Suppose we assume an average delivery distance of 10 miles for the demand that is collocated with our DC's. That comes to $\frac{1}{4} \times 25,000,000$ CWT $\times 10$ miles $\times$ \$0.0075(CWT-mi) $=$ \$468,750. For the non-collocated demand, suppose that the average delivery distance is 156 miles [453, Figure 5]. That comes to $\frac{3}{4} \times 25,000,000$ CWT $\times$ 156 miles $\times$ \$0.0075(CWT-mi) $=$ \$21,937,500. Total predicted outbound transportation cost is the sum of these two numbers, or $22,406,250. This prediction is $6,093,750 below the adjusted figure of $28,500,000 given at the conclusion of the discussion of Q1. This suggests that the average delivery distance must actually be considerably greater than 156 miles, perhaps as much as 200 miles, which in turn suggests that our current DC's are not well situated with respect to current concentrations of demand.

To be conservative, I will assume that improved DC locations could achieve half of the savings potential indicated above, namely $3 million. This may, however, be gained at the price of increased inbound freight cost. Inbound costs should rise much less than outbound costs will fall. One reason is that our plants are scattered all over the country, so any DC move **away** from one plant will tend to be **toward** another plant. Cost changes thus will tend to cancel out. Another reason is that empirical studies with simplified one plant models show that inbound transportation cost tends to change very little with the number and location of DC's provided there are at least 10 of them ([453, equation (14)]). Thus, I believe that no more than half of the outbound cost savings would be lost in increased inbound costs. This leads to the following revision of the cost breakdown given at the conclusion of the discussion of Q1:

DC fixed costs	$ 6,150,000
DC variable costs	18,450,000
Inbound transportation	43,500,000
Outbound transportation	25,500,000
	$93,600,000

Q3. Do we have too many or too few DC's?

I believe that we have too few DC's. My reasons are several. First of all, demand growth implies that new DC's should be added over time because there is greater volume over which to spread the DC fixed costs. Second, more DC's are needed when outbound freight rates inflate faster than inbound freight rates because more DC's reduce the amount of outbound shipping. We are behind the times on both counts, as we have had 13 DC's for 4 years. Another tipoff that we have too few DC's is the fact that total outbound freight cost is more than 4 times total DC fixed cost, even after the adjustments made in my discussion of Q1 and Q2. An idealized analysis suggests that this ratio should be more like 2 than 4, and hence that we need to increase the number of DC's (to decrease the numerator and increase the denominator of this ratio). The idealized analysis I have in mind is as follows:

If a DC has fixed cost f $\$/yr$, lies in a part of the country having demand evenly distributed with density ρ CWT/(mi^2-yr), can deliver for t \$/(CWT-mi), and has a circular service area of S square miles, then it is an elementary exercise to show that the total associated annual cost is $f + .377\sqrt{S}(\rho S)t$. Total annual cost per CWT throughput (divide by ρS) is minimized when $S^* = 3.05(\rho t/f)^{-2/3}$. Note that the outbound transportation cost at S^* is $.377\rho[3.05(\rho t/f)^{-2/3}]^{3/2}t = 2f$.

This analysis can be applied to our problem with these numbers: $f = \$6,150,000/13 = \$473,077$; $\rho = 25,000,000$ CWT$/2,000,000$mi^2 = 12.5 CWT/mi^2 (2 million square miles is actually $\frac{2}{3}$ of the land area of the continental U.S., to account for the absence of any significant population in much of the western central part of the country); $t = \$0.0075$(CWT-mi). The result is $S^* = 89,730mi^2$. To cover 2 million mi^2, this requires **22 DC's**.

This estimate of 22 for the optimal number of DC's, besides being based on extremely simplified assumptions, does not consider the influence of inbound transportation costs or the fact that DC's tend to coincide with major demand concentrations. These considerations introduce cost tradeoffs that tend both to increase and decrease the optimal number of DC's. I will not pursue such refinements here.

Suppose for the sake of argument that 22 is the best number of DC's for us. Let's examine the cost consequences. Total fixed cost should go up by about $\$473,077 \times (22-13) = \$4,257,700$. Total DC variable cost should not change significantly, as the variation between individual DC variable cost rates is not great. Total inbound cost should not change much, for reasons described under Q2. I'll assume a generous $1 million increase. Total outbound cost should decrease significantly. The idealized analysis discussed above predicts that total outbound cost varies as one over the square root of the number of DC's (See, for example, [451, equation (3)]). It follows that total outbound cost will go down to a fraction of $\frac{1/\sqrt{22}}{1/\sqrt{13}} = 0.7687$ of the former level. But this neglects the fact that many of the 9 new DC's will fall on major demand concentrations, thereby reducing outbound freight even more. Suppose that the 9 new DC's are collocated even half as well as the current 13. Then an additional $\frac{1}{2} \times \frac{9}{13} \times .25 = 9.375\%$ of all demand will be collocated, which will reduce the 0.7686 factor to $(1-0.09375) \times 0.7686 = 0.6965$. Applying this to the $25,500,000 figure given at the end of the discussion on Q2, we obtain a savings of $7,740,000. This is enough to overcome the increases in fixed and inbound costs, with a net annual savings of $2,480,000. With these assumptions, the cost breakdown given at the end of the discussion of Q2 becomes:

DC fixed costs	$10,400,000
DC variable costs	18,450,000
Inbound transportation	44,500,000
Outbound transportation	17,761,000
	$91,110,000

Summary

I have argued that possible improvements in our current distribution system can be viewed as occurring in three phases corresponding to the three questions stated at the outset. The cost breakdown is summarized as follows (costs in millions):

	Current	(Q1) Better Utilization	(Q2) Relocate Current DC's	(Q3) Add DC's
DC fixed costs	$ 6.150	$ 6.150	$ 6.150	$10.400
DC variable costs	18.500	18.450	18.450	18.450
Inbound transp	42.850	42.000	43.500	44.500
Outbound transp	29.100	28.500	25.500	17.760
Total	$96.600	$95.100	$93.600	$91.110
Savings over previous system		$1.500	$1.500	$2.490

Adding more DC's (about 9 of them) will make marketing happy, as each new DC would increase our "market presence" in a new area and probably lead to improved market penetration. Notice also that adding DC's should greatly reduce outbound average shipping distances, which tends to reduce our vulnerability to continued rapid increases in motor carrier rates.

The total potential savings of about 5.5%, an estimate that I view as quite conservative, falls within the 5 – 15% range which distribution consultants claim is typical for a full scale modeling study.

I hope that these "back of the envelope" arguments will suffice to help you decide whether to sponsor a modeling effort. Comparable studies at other firms suggests that the costs of such a study would be under $100,000 and take about 6 months to complete. It looks to me like a high payoff venture.

[memo 3]

TO: Director of Management Science
FROM: Vice President of Logistics
DATE: July 1

Your response to my June 1 memo makes a good case for undertaking a thorough analysis. With it, I can justify making a substantial commitment in labor and resources. Let's go ahead with the project. The main target issues I would like to address are:

- How many DC's should we have?
- Where should the DC's be located?
- What size should each DC be?
- How should our 7 plants be loaded, and how should their output be allocated to the DC's?
- Which customers should be assigned to each DC?

Although cost minimization is the basic objective, customer service should also be considered. The physical proximity of DC's to their markets is an important factor.

Please conduct the analysis so as to continue our present policy of assigning each customer to a single private full line DC.

One more thing. My last employer had a real disaster with a computer-based model for inventory management. Finished goods inventory went up and up and up some more instead of down as promised. That experience taught me that it is important for executives to keep in close touch with their fancy models. Consequently, I will ask that you explain to me in non-mathematical terms the model as it evolves, and whatever results come out of the model after it is built. I want you to help me discover not only *what* to do, but also *why* to do it.

[memo 4]

TO: Vice President of Logistics
FROM: Director of Management Science
DATE: October 1
RE: Distribution Planning Model Synopsis

The data development stage of the project is virtually complete now. Soon we will be commencing verification and validation. You mentioned in your July 1 memo your desire to "keep in close touch" with the model and to understand why, as well as merely know what it is trying to tell us. I'm glad that you feel this way, as it coincides exactly with my own philosophy.

An important step is for you to have at your fingertips a clear and concise synopsis of the model. Such a synopsis is attached, together with a report on the supporting data. [The model structure is included here, but for brevity the detailed data are omitted.] After scrutinizing it for possible errors, you can keep it at hand for ready reference as we move into the formal analysis stage of the project.

THUMBNAIL MODEL SKETCH and DATA ELEMENT CHECKLIST

7 PLANTS ⟹ 35 CANDIDATE DC's ⟹ 150 CUSTOMER ZONES

(1) List of product groups (10)

(2) List of plants (7)

(3) Plant capacities in CWT /yr (plant × product)

(4) Unit production cost in $/CWT (plant × product)

(5) List of candidate sites for DC's (35); all private and full line

(6) Lower limit on each DC throughput is 100,000 CWT/yr

(7) Fixed cost for each DC in $/yr

(8) Variable cost in $/CWT by DC (same for all products)

(9) List of customer zones (150 most populous S&MM metropolitan markets)

(10) Annual customer demand in CWT/yr (customer zone × product)

(11) Single sourcing rule: each customer zone assigned uniquely

(12) Net selling price (customer zone × product); omitted

(13) List of permissible inbound links and freight rates in $/CWT (same for all products, nearly all rail)

(14) All possible outbound links up to 1500 miles long included (a total of 3728 DC/customer pairs); freight charge is $0.0075/(CWT-mi) for any product (nearly all truck)

[memo 5]

TO: Vice President of Logistics
FROM: Director of Management Science
DATE: December 10
RE: Formal Analysis: Summary of Key Findings

Since completing the verification and validation exercises about a month ago, we have as you know been busy using the model as a tool to study how our distribution system can be improved. Your active participation in these studies has been invaluable. Although you have seen most of the results already in raw form, we are now in a position to summarize the key findings in a coherent way.

I want to stress that nothing in this memo is intended as a recommendation. We are, however, at the point where the findings and insights summarized here must be interpreted in light of various considerations outside the scope of the formal analysis. The result of that exercise will be specific, prioritized recommendations for action.

This memo makes three passes at summarizing the key findings. The first establishes that current annual distribution costs can (according to the model) be reduced from $96,600,000 to $88,810,000. This represents a projected savings of $7,790,000. The second pass establishes that the projected savings can be attributed to three distinct types of improvement carried out sequentially:

Types of Improvement	Savings
Improve plant loading, product allocation to DC's, and customer assignments (keeping the current DC's)	$2,269,000
Improve DC locations (keeping the number at 13)	2,412,000
Move to a network with 21 DC's	3,109,000
	$7,790,000

The third pass attempts to achieve a deeper understanding of each of the three types of improvement.

The figures mentioned above are exclusive of production costs. You will recall

4-10, owing to the severe limitations of our cost accounting systems. It turns out that the unit production costs we did include for products $1-3$ exert only a negligible influence on the target issues of interest. Hence this memo intentionally disregards the role of production costs.

FIRST PASS: The Bottom Line

The bottom line is that projected annual savings amount to $7,790,000. Here is a comparison of current annual costs (in millions) with what is projected under the least cost distribution network, which has 8 more DC's:

	Current	Least Cost Network	Savings (Loss)	% Change from Current
DC fixed costs	$ 6.150	$ 9.737	($3.587)	58
DC variable costs	18.500	18.727	(0.227)	1
Inbound transp	42.850	43.428	(0.578)	1
Outbound transp	29.100	16.918	12.182	-42
Total	$96.600	$88.810	$7.790	-8

Considering our corporate challenge to reduce logistics costs, I think you will be pleased to see that massive savings in outbound trucking materialized as predicted. These more than offset increases in other cost categories caused (mainly) by the addition of new DC's. Profits may well go up more than $7,790,000 because our analysis assumed demand to be fixed, whereas it should actually increase somewhat owing to the addition of 8 DC's. Marketing claims that each new DC improves local market share through improved "market presence." One measure of market presence is average delivery distance, which decreases from about 150 miles with the current network to about 90 miles with the least cost network.

The answers to all of the target issues listed in your July 1 memo are available in a detailed set of reports describing the least cost network.

SECOND PASS: Three Steps to a Least Cost Network

The cost comparison given above shows how the projected $7,790,000 savings is distributed by cost category, but does not tell us why the cost categories change as they do in going from the current to the least cost network. My aim now is to identify the major reasons why the costs change as they do.

The explanation has to do with the different kinds of changes which might be made in our current distribution network. There are 5 possible kinds of changes:

A. change plant loadings (how much of each product is made at each plant);
B. change product allocations (the shipping pattern from plants to DC's);
C. change the assignment of customer zones to DC's;
D. change the DC locations (keeping the number the same);
E. change the number of DC's.

If we allow no changes, the Distribution Planning Model will simply mimic (simulate) our current distribution network and will therefore yield a total cost of $96,600,000. As more and more freedom is allowed to make changes, the model will be able to drive the total cost closer and closer to the $88,810,000 floor.

It is instructive to examine what happens when the model is given only partial freedom to make changes. In this way we can better understand where the ultimate total savings come from. We have examined two such cases. These, along with the current and "full freedom to change" cases, can be portrayed as follows:

```
                A          B          C          D          E
"Current"   ================fixed==================
"Step 1"    *****optimize******************=======fixed=====
"Step 2"    *****************optimize*****************===fixed==
"Step 3"    **********************optimize**********************
```

The model also lends itself to the case where optimization occurs only over A and B, but we did not pursue this one. The managerial interpretation of these cases is clear. Step 1 asks "How much improvement is possible by better utilization of the current 13 DC locations?" Step 2 asks "What is the least cost distribution network having 13 DC's?" Step 3 asks "What is the least cost distribution network?"

The <u>differences</u> are particularly revealing, as they enable the total $7,790,000 savings to be decomposed into components associated with utilization alone, locational choice, and a change in the number of DC's. These differences and the cases themselves are summarized below. You may wish to make a comparison with a similar summary at the end of my June 10 memo.

This summary reveals that improved utilization of the current 13 DC's could save $2,269,000, improved locations of the 13 DC's could save another $2,412,000, and going to a 21 DC network could save $3,109,000 more.

THREE STEPS TO THE LEAST COST DISTRIBUTION NETWORK

		— Step 1 —			— Step 2 —			— Step 3 —		
Cost Element	Current Network	Optimal Util.	Savings (Loss)	%Δ	Optimal Loca's	Savings (Loss)	%Δ	Optimal Number	Savings (Loss)	%Δ
DC fixed costs	$ 6.150	$ 6.139	$0.011	-0.2	$ 6.000	$ 0.139	-2.3	$ 9.737	($3.737)	62.3
DC var. costs	18.500	18.462	0.038	-0.2	18.716	(0.254)	1.4	18.727	(0.011)	0.1
Inbound tran.	42.850	41.777	1.073	-2.5	42.693	(0.916)	2.2	43.428	(0.735)	-1.7
Outbound tran.	29.100	27.953	1.147	-3.9	24.510	3.443	-12.3	16.918	7.592	-31.0
Total	$96.600	$94.331	$2.269	-2.3	$91.919	$2.412	-2.6	$88.810	$3.109	-3.4

Legend:

1. All costs in millions.

2. Savings (loss) and % are with reference to the preceding step.

______________[end of memo 5]_________________________

9.3 Diagnostic Tools

Managers are unlikely to accept analysis they do not satisfactorily understand. Mathematical programming models can deliver an optimal solution for a given set of input data, but they do not explain *why* the solution is what it is. That is why our case study took pains to prepare the way for, and then to interpret, the optimal results by intuitively reasonable explanations. In this section we consider diagnostic tools that, among other functions, promote understanding of the numerical results.

Diagnostic tools are methods suitable for quickly detecting high potential opportunity areas worthy of further study. Given the complexity of the real-world, these tools tend to be based on stripped-down models that attempt to capture the essence of a problem with minimal data requirements assumptions). As such, diagnostic tools serve to build and complement intuition. Both the importance and the lack of such methods in the literature are noted in [451, 453, 500], [566, chapter 7], [1014, see section 2.6.2 in particular].

Memo 2 illustrates the use of a simple model for both diagnosis and developing insight into the nature of cost trade-offs. In the remainder of this section we briefly present a generalization of this model, known in the operations research literature as the general optimal market area (GOMA) model, and comment on its potential range of application. We begin with the assumptions.

1. Demand per year is evenly distributed over the distribution territory with density ρ per square mile.
2. Given market area S served by a facility, the annual facility volume is ρS and the facility operating cost is $f(\rho S)^\alpha + v\rho S$. Economies of scale are reflected in $f(\rho S)^\alpha$, where $0 \leq \alpha < 1$, with smaller values of α implying greater relative economies of scale.
3. Unit transportation cost as a function of miles traveled, δ, is $t\delta^\beta$. Economies of distance exist when $\beta < 1$. If $\beta = 1$, then t represents the unit transportation cost per mile.
4. Various regular two-dimensional market shapes may be specified (e.g., hexagon, circle, diamond).
5. Various distance norms may apply (e.g., Euclidean, rectilinear).
6. The objective is to minimize cost.

For many combinations of β, market shape, and distance norm, transportation cost is proportional to the market area raised to the $\frac{\beta}{2}$ power; that is, average transportation cost per unit $= t\sigma S^{\frac{\beta}{2}}$ where σ is the *configuration factor*. The value of σ depends on β, the market shape, and the distance norm. For example, with a circular market shape and the Euclidean distance norm, $\sigma = \frac{2}{(2+\beta)\pi^{\frac{\beta}{2}}}$. See [391] for other values of σ. When $\beta = 1$, σ is simply the average distance from the market center to customers for a one square mile market area.

Based on the above information, we may express the average facility operation and transportation cost per unit as

$$f(\rho S)^{\alpha-1} + v + t\sigma S^{\frac{\beta}{2}}, \tag{9.1}$$

which is minimized at $$S^* = \left\{ \frac{2(1-\alpha)f}{\beta t\sigma} \rho^{\alpha-1} \right\}^{\frac{2}{\beta+2(1-\alpha)}}. \tag{9.2}$$

The ratio of transportation cost to fixed facility cost at S^* is

$$\frac{t\sigma(S^*)^{\frac{\beta}{2}}}{f(\rho S^*)^{\alpha-1}} = 2\frac{1-\alpha}{\beta} \, . \tag{9.3}$$

Expression (9.1) gives an indication of the cost structure present in distribution system design problems. As illustrated in memo 2, expressions (9.2) and (9.3) are helpful for evaluating the sensitivity of an efficient design to changes in demand and cost rates, and for quickly assessing the potential value of redesigning an existing system.

We have limited our discussion to a basic form of the GOMA model. Other features that can be incorporated into the model include inbound transportation costs [453], demand that is dependent upon distance from the facility [391], a network comprising primary and secondary facilities [1182], and demand that changes over time [1183].

9.4 Algebraic Language Tools

Algebraic modeling languages interfaced with standard optimizers provide an intermediate tool for quickly and easily developing prototype mathematical programming representations of distribution system design models. Examples include AMPL [418], GAMS [148], and LINGO [1050]. An algebraic programming language overcomes some of the awkwardness that attends standard commercial general-purpose MIP software. Such software usually requires coding front and back end software for its use. Changes in the parameter values for sensitivity analysis are often difficult and changes to the model structure can require excessive labor. Porting to new computer environments poses additional challenges. It is important to stress, however, that specialized algorithms will still be needed to solve large-scale distribution system design problems.

We use LINGO in the following illustration of algebraic modeling languages, but there is no intention to favor any particular software product. It is applied to a model described by Pooley [966] for a study of facilities at Ault Foods Limited. Super-LINDO was used as a solver. (LINDO is the background solver for LINGO.) The model, in which both plants and warehouses are located is given by:

Minimize: $\sum_j [e_j a_j + \sum_{ik} p_{ij} x_{ijk}] + \sum_k [f_k z_k + v_k \sum_{il} D_{il} y_{kl}]$

$$+ \sum_{ijk} c_{ijk} x_{ijk} + \sum_{ikl} t_{ikl} D_{il} y_{kl}$$

subject to:

production capacity (for all j): $\sum_{ik} x_{ijk} \leq P_j a_j$

warehouse capacity (for all k): $\sum_{il} D_{il} y_{kl} \leq W_k z_k$

production meets demand (for all $i$and k): $\sum_j x_{ijk} = \sum_l D_{il} y_{kl}$

single-sourcing (for all l): $\sum_k y_{kl} = 1$

with $x_{ijk} \geq 0$; a_j, z_k, and $y_{kl} = 0, or 1$;

where:

- i index for commodities;
- j index for plant sites;
- k index for warehouse sites;
- l index for customer zones;
- e_j fixed portion of possession and operating costs for plant j;
- f_k fixed cost for warehouse k;
- v_k average unit variable cost of throughput for warehouse k;
- p_{ij} average unit cost of producing commodity i at plant j;
- c_{ijk} average unit cost of shipping commodity i from plant j to warehouse k;
- t_{ikl} average unit cost of shipping commodity i from warehouse k to customer zone l;
- D_{il} demand for commodity i in demand zone l;
- P_j production capacity of plant j;
- x_{ijk} amount of commodity i shipped from plant j to warehouse k;
- a_j zero-one variable equal to 1 if a plant is established at site j, and equal to 0 otherwise;
- z_k zero-one variable equal to 1 if a warehouse is established at site k, and equal to 0 otherwise;
- y_{kl} zero-one variable equal to 1 if warehouse k is assigned to demand zone l, and equal to 0 otherwise;

and all parameters measured on an annual basis. Solution times were 15-90 minutes using a 386 PC. There were 6 commodity groups, 10 plant sites, 13 warehouse sites, and 48 customer zones. The distribution network had a sparse structure as viable outbound links were determined through discussions with marketing representatives to satisfy service constraints, and not all possible plant-warehouse combinations were included [967].
We developed the following LINGO expression of the *dense network* form of the model (with data based on the Huntco Foods case in [1210]):

```
MODEL:
SETS:
PLANTS/1..3/:PCAP,FCP,PSITE;
WARE/1..5/:WCAP,FCW,WCOST,WSITE;
CUST/1..6/;
ITEMS/1..2/;
INBOUND(ITEMS,PLANTS,WARE):INCOST,VOL;
OUTBOUND(ITEMS,WARE,CUST):OUTCOST;
COMBIJ(ITEMS,PLANTS):PCOST;
```

```
COMBIK(ITEMS,WARE);
COMBKL(WARE,CUST):YWC;
COMBIL(ITEMS,CUST):DEM;
ENDSETS
DATA:
FCP = 1,1.1,1.5;
FCW = .899,.899,.899,.899,.899;
WCOST = .6,.5,.4,.7,.9;
PCAP = 15,12,13;
WCAP = 9,9,9,9,9;
DEM = 2,2.5,4,2,3.5,5,
     1,2.5,3.5,3,4.5,2;
INCOST = .20,.57,.58,.53,.69,
     1.04,.91,.93,.91,.80,
     .59,.34,.47,.31,.32,
     .85,.76,.67,.58,.49,
     .29,.38,.47,.56,.65,
     .15,.16,.17,.18,.19;
OUTCOST = .34,.63,.48,.82,.76,1.07,
     .62,.46,.33,.63,.61,.92,
     .64,.39,.53,.75,.56,1,
     .6,.42,.33,.65,.59,.95,
     .75,.42,.55,.49,.44,.8,
     .25,.35,.45,.55,.65,.75,
     .98,.88,.78,.68,.58,.48,
     .15,.25,.35,.45,.55,.65,
     .91,.81,.71,.61,.51,.41,
     1.1,1.2,.89,.65,.75,.34;
PCOST = 2,4,3,
     5,3,2;
ENDDATA

! The objective function;
[COST] MIN =
@SUM( INBOUND(I,J,K): INCOST(I,J,K)*VOL(I,J,K))
+ @SUM( OUTBOUND(I,K,L): OUTCOST(I,K,L)*DEM(I,L)*YWC(K,L))
+ @SUM( PLANTS(J): FCP(J)*PSITE(J)
+ @SUM( COMBIK(I,K): PCOST(I,J)*VOL(I,J,K)))
+ @SUM( WARE(K): FCW(K)*WSITE(K)
+ WCOST(K)*@SUM( COMBIL(I,L):YWC(K,L)*DEM(I,L)));
! Production capacity constraints;
@FOR( PLANTS(J): [PCAP]
@SUM( COMBIK(I,K): VOL(I,J,K)) < PCAP(J)*PSITE(J));

! Production meets demand;
@FOR( ITEMS(I):
@FOR(WARE(K): [BALANCE]
@SUM( PLANTS(J):
VOL(I,J,K)) - @SUM( COMBKL(K,L) DEM(I,L)*YWC(K,L)) = 0))
```

```
! Warehouse capacity constraints;
@FOR( WARE(K): [WCAP]
@SUM( COMBIL(I,L): DEM(I,L)*YWC(K,L)) < WCAP(K)*WSITE(K));
! Single-sourcing constraints;
@FOR( CUST(L): [SOURCING]
@SUM( WARE(K): YWC(K,L)) = 1);
@FOR( COMBKL(K,L): [FORCE] YWC(K,L)<WSITE(K));
! Set the binary restrictions;
@FOR( WARE(K):
@BIN(WSITE(K)));
@FOR( COMBKL(K,L):
@BIN(YWC(K,L)));
@FOR( PLANTS(J):
@BIN(PSITE(J)));
END
```

The @SUM operator replaces the $\sum$, and the @FOR defines loops through the indices. The bracketed terms in the model are labels for ease of identification of solution report details. The forcing constraints $y_{kl} \leq z_k$ have been added here to tighten the formulation. They render the LP relaxation "integer friendly," and can speed up the MIP solution process. ReVelle [994] reviews examples of such models and the history of such tightening constraints (see also [242, 454, 795, 869]). The small numerical example resulted in premature termination of the solver, while looser plant and warehouse capacity constraints allowed quick solution times using a 486 PC running at 25 mhz, and a student version of LINGO. (The model could be converted to MPS format using the SMPS command if needed for use with another solver. Also, sparse link sets could be used to model sparse network structures, and an ϵ-optimal termination criterion could be implemented.)

The main limitation of algebraic modeling languages for distribution system design is that they have not been interfaced with the highly specialized solvers needed to optimize at the level of detail needed in most real applications. When this limitation is overcome, such languages will graduate from a supplementary role to a primary one.

9.5 Conclusions

We began with a series of memos in order to illustrate analysis of a distribution system in a way that is compatible with developing intuition into the problem. We then briefly described two types of distribution system design tools, diagnostic tools and language tools. The strength of diagnostic tools is in quickly assessing the pay-off potential of change and the consequent merit of more detailed study. We see diagnostic tools as complementary to the specialized optimizers required for most practical settings. The algebraic language tools provide efficient model specification and, at the least, a first step in more detailed analysis of a company's logistics system. Applications of these tools alone may even be sufficient in some cases to establish the confidence needed for a final decision. We have purposely left our descriptions concise and rely on the following annotated bibliography as a reference for further detail.

9.6 Annotated Bibliography

9.6.1 Related to Diagnostic Tools

Two papers that are helpful for providing a general background on the use of diagnostic tools for analysis are [889, 518]. Both authors cover a number of applications and give examples of approximate models for quick analysis.

The literature on market area models is extensive and spans many disciplines. The origins of modern activity on market area models can be traced back at least 50 years to the work of German economist August Lösch (1906 - 1945). An excellent review of the literature is given in [391]. We will limit overlap by restricting mention to a few papers that help illustrate the range of application of market area models. We will also point out more recent work and identify sources for average distance formulas.

The papers [744, 451, 453] give more background on the diagnostic tool employed in memo 2. Illustrations in [453] are drawn from applications in the auto parts, consumer products, food, and mining industries. Smith [1089] describes market area models for estimating the number of service personnel who are responsible for machine repairs (e.g., photocopiers, computer systems) in assigned geographic territories. The tools are relevant for planning social services as well (public health nurses, social workers). Other public sector applications include capacity and market area decisions for solid waste transfer-stations [1212], bus garages [1175], and fire stations [1205]. In economic planning and analysis, market area models have been used to estimate the benefits due to transportation improvements [861] and are well-established in analyzing spatial pricing policies under competition [72].

Examples of recent work on market area models for distribution system design include [158, 159, 160, 161, 162, 163, 1182, 1183]. Campbell [158, 159, 161, 162] analyzes a network comprising a single plant supplying distribution centers which serve local markets. Hub location is addressed in [163] and the merits of a two-echelon distribution system with primary and secondary facilities are analyzed in [1183]. Dynamic demand is introduced in [160] for the case of deterministic increasing demand over time, and in

[1183] for a system with changing and uncertain demand.

Market area models make use of average distance expressions (see [366, 740, 1101] for formulas). Average distance measures are used to study the role of transshipment centers in many-to-many logistic networks in which vehicles make deliveries from many sources to many destinations [256]. Near optimal network geometry is modeled using a small set of parameters.

9.6.2 *Related to Algorithmic Tools*

Aikens [13] reviews operations research contributions to the modeling and analysis of distribution systems design that involves "determination of the best sites for intermediate stocking points." Models are classified according to:

- whether the underlying distribution network (arcs and/or nodes) is capacitated;
- the number of warehouse echelons (zero, single, or multiple);
- the number of commodities (single or multiple);
- the underlying cost structure for arcs and/or nodes (linear or nonlinear);
- whether the planning horizon is static or dynamic;
- patterns of demand (e.g., deterministic or stochastic, influence of location, etc.);
- the ability to accommodate side constraints (e.g., single-sourcing, choice of only one from a candidate subset, etc.).

Discussion begins with heuristic and branch-and-bound approaches applied to the single-commodity, uncapacitated, zero warehouse echelons, linear cost model often referred to as the simple plant location model, and ends with Benders decomposition applied to the multicommodity, capacitated, single warehouse echelon case as proposed in [454], and elaborated upon in [455]. (A numerical example of the Benders decomposition approach appears in [786]. Early recognition of the potential of Benders decomposition is discussed by Balinski [50] who cites the working paper [49].)

Work considering stochastic systems is sparse, and work directly considering nonlinear costs is practically nonexistent. Of course piece-wise linear approximations for nonlinear concave throughput cost functions were proposed long ago, as in [365]. Geoffrion [452] provides guidance on constructing optimal objective function approximations. A recent contribution considering the stochastic demand, zero warehouse echelons, capacitated, single-commodity case is [736], "for which no exact algorithm had been previously developed." A brief literature review is provided (and [771] is suggested for further background). Solution of large-scale stochastic demand problems involving multiple commodities, intermediate facilities, with various capacity constraints awaits further developments.

Beasley [66] reviews the literature of capacitated, single commodity warehouse location problems (that omit the plant-warehouse linkage), and reports on a tree search procedure capable of solving up to 500 potential warehouse locations and 1000 customer zones. Use of Lagrangian relaxation for bounding and imbedding problem reduction rules follows on a rich history of such strategies in solving facility location problems.

Structure-exploiting factorization has also contributed to the evolution of algorithms applied to logistics models as discussed in [458]. Examples of advances in factorizing embedded network structures are [151, 796], while [150] provides an example of the robustness of the Benders decomposition strategy. The latter work introduces "goal cuts" within a decomposition approach to accelerate convergence when solving multicommodity production/distribution problems for Nabisco Brands, Inc. Plant location decisions are combined with assignment of production facility types to the plants.

Most published accounts of optimizing models applied to distribution system design have been of the MIP variety where binary facility location variables are used to select from a finite set of candidate sites. In [447] a continuous model (location can take place anywhere in the plane) is paired with the standard MIP approach to decide supply points for shipping to customers of a large Belgian brewery. Tactical issues involved in implementation are addressed. A good account of the practical issues involved in successfully applying MIP models to support both long-term and short-term distribution system planning decisions is given in [483]. The development and use of a decision support system (DSS) for DowBrands, Inc., based on an MIP model of a distribution system having two warehouse echelons is described in [1011]. Geoffrion and Powers [457] discuss the need for comprehensiveness in the distribution planning system. Shapiro, Singhal, and Wagner [1064] use a case study to illustrate how integrated logistics planning advocated by logistics practitioners (and espoused in [968]) can be successfully employed through DSSs based on mathematical programming models.