Jennifer Shang, Tuba Pinar Yildirim, Pandu Tadikamalla, Vikas Mittal, & Lawrence H. Brown

Distribution Network Redesign for Marketing Competitiveness

This article reports on a marketing initiative at a pharmaceutical company to redesign its distribution network. Distribution affects a firm's cost and customer satisfaction and drives profitability. Using a nonlinear mixed-integer programming model, the authors develop a distribution network with a dual emphasis on minimizing the total distribution costs and improving the customer service levels. Specifically, they address the following issues: They (1) determine the optimal number of regional distribution centers the firm should operate with, (2) identify where in the United States the firm should locate these distribution centers, (3) allocate each retailer/customer distribution center to an appropriate regional distribution center, and (4) determine the total transportation costs and service level for each case. Finally, they conduct a sensitivity analysis to determine the impact of changes in problem parameters on the optimality of the proposed model. This marketing initiative at the studied firm reduced the total distribution costs by $1.99 million (6%) per year, while increasing the customer on-time delivery from 61.41% to 86.2%, an improvement of 40.4%.

Keywords: location theory, optimization, decision calculus, network design, supply chain model

As the service element of customer experiences becomes more important, location and convenience have emerged as major factors in consumer decisions for products and services (e.g., Chan, Padmanabhan, and Seetharaman 2007; Devlin and Gerrard 2004; Ghosh and Craig 1986; Mulhern 1997; Thelen and Woodside 1997). As Ghosh and Craig (1983, p. 56) argue,

> A good location provides the firm with strategic advantages that competition may find difficult to overcome. While other marketing mix elements may be easily changed in response to a changing environment, store locations represent long-term investments that can be changed only at a considerable cost.

Over the past decade, retailers have attempted to manage their store locations strategically to reach more and more consumers (Langston, Clarke, and Clarke 1997). As the number of retail outlets has increased, manufacturers have responded by modifying their distribution network to eliminate stockouts, minimize late deliveries, and reduce supply costs by changing shipping routes, relocating distribution centers, and reconfiguring warehouses (Deveci-Kocakoc and Sen 2006). Within the marketing discipline, although research on retail locations has been conducted by many scholars (e.g., Ghosh and Craig 1983, 1986; Kuo, Chi, and Kao 2002; Mahajan, Sharma, and Kerin 1988; Pinkse, Slade, and Brett 2002), there is a dearth of research on how manufacturers can design their distribution network in response to retailer location networks. Korpela and Lehmusvaara (1999) empirically study retail clients of a manufacturer and find that factors such as delivery time, quality, total cost, and ability to meet consumers' urgent/special needs are key drivers that affect retailers' decisions to carry the manufacturer's products. These factors should guide decisions about distribution centers and warehousing locations.

Rust and colleagues (2004) highlight the importance of efficiency in marketing systems to make marketing more financially accountable to top management. Using the nonlinear mixed-integer programming approach, we develop a distribution network that not only improves efficiency by minimizing the total distribution costs but also improves customer service levels. We illustrate our approach in the context of a global pharmaceutical firm in which the marketing department, responding to the needs of the retailers and top management, launched an initiative aimed at reengineering the firm's distribution network.

The model we propose herein addresses the following issues: (1) determining the "optimal" number of regional distribution centers (RDCs) the manufacturing firm should operate with, (2) identifying where in the United States the firm should locate these distribution centers, (3) allocating each of the retailer/customer distribution centers (CDCs) to an appropriate warehouse, and (4) determining the total transportation costs and service level for the optimal scenario, as well as other scenarios. A sensitivity analysis examines the impact of changes in model parameters on the optimality of the proposed model. Finally, to understand

Jennifer Shang is Associate Professor of Business Administration (e-mail: shang@katz.pitt.edu), Tuba Pinar Yildirim is a doctoral candidate (e-mail: tyildirim@katz.pitt.edu), and Pandu Tadikamalla is Professor of Business Administration (e-mail: pandu@katz.pitt.edu), Katz Graduate School of Business, University of Pittsburgh. Vikas Mittal is J. Hugh Liedtke Professor of Marketing, Jones Graduate School of Management, Rice University (e-mail: vmittal@rice.edu). Lawrence Brown is Vice President of Customer Supply and Planning, GlaxoSmithKline Company (e-mail: Lawrence.H.Brown@gsk.com).

ISSN: 0022-2429 (print), 1547-7185 (electronic)

the quality of our methodology, we compare the solution approach with two other heuristics. This marketing initiative was able to reduce the total distribution costs of the studied firm (GlaxoSmithKline [GSK]) by $1.99 million (6%) per year and to increase on-time delivery from 61.41% to 86.2%, an improvement of 40.4%.

Answering the call of Rust and colleagues (2004, p. 84) that marketing's contribution should specifically focus on "core business processes and efficient supply chain processes," we show how to improve the efficiency and effectiveness of distribution systems simultaneously. In doing so, we show the interdependency of retail strategy and distribution strategy in terms of location analysis. Despite a large focus on retail stores' location selections (e.g., Devlin and Gerrard 2004; Ghosh and Craig 1986; Mulhern 1997), their interdependence on the distribution network has not been examined. Yet successful retailers, such as Wal-Mart and Target, demonstrate the need for incorporating distribution strategy for marketing success. Finally, we show how input from marketing managers can be gainfully used in distribution design. This not only illustrates the concept of decision calculus (Chakravarti, Mitchell, and Staelin 1979; Little 2004) but also shows how the decision calculus approach developed in marketing can be applied more broadly. For example, we obtained managerial input on problem formulation, along with the various parameters of the distribution network.

We organize the rest of the article as follows: We begin with a review of the distribution network design from marketing and operations management literature regarding the structure of a firm's supply chain. We provide details of the generic network design model and its analysis in the "Problem Formulation and Methodology" section. Then, we apply the generic model to GSK. We follow this with the results of the application, conduct a sensitivity analysis of the recommended solution methodology, and compare the performance of our approach with other heuristics. Finally, we provide a summary of findings and note the limitations of this research.

Literature Review

Motivated by the importance of store location and customer convenience as key elements of marketing strategy, scholars have developed models to guide optimal location decisions for retailers and service providers (Bucklin 1967; Cox 1959; Ghosh and McLafferty 1982; Mulhern 1997). Early models used regression analysis to determine store locations (Lord and Lynds 1981), while later models also incorporated insights from game theory and decision theory (e.g., Davis 2006; Ghosh and Craig 1983, 1986). More recently, Chan, Padmanabhan, and Seetharaman (2007) estimated an econometric model that incorporates the geographic location of retailers and models the price competition among them to determine consumer policy implications. Empirically, models have been developed to incorporate the spatial variability in customer tastes when determining store locations (Donthu and Rust 1989; Mittal, Kamakura, and Govind 2004; Rust and Donthu 1995). Thus, in marketing, there is a rich tradition of examining retail locations from the retailer's or customer's perspective. However, the location choices made by manufacturers to support retailer networks have received relatively little attention in the marketing literature. Yet it is well known that manufacturer decisions can have a critical effect on the marketing success of downstream retailer partners (Iyer and Bergen 1997; Kadiyali, Chintagunta, and Vilcassim 2000; Murry and Heide 1998). This is the focus of the current research.

The field of location analysis has been extensively studied (for a review, see Brandeau and Chiu 1989; Daskin 1995). The location and allocation decisions in supply chain network design, including the choice of the number, site, and capacity of facilities, as well as assigning customers to these facilities, have significant long-term impacts on the efficiency of the network. Model formulations and solution algorithms that address these issues vary widely in terms of fundamental assumptions, mathematical complexity, and computational performance. We review key developments in this literature.

Design of a Distribution Network

Research in location–allocation often focuses on cost reduction, demand capture, equitable service supply, and fast response time. Baumol and Wolfe (1958) were the first to describe a distribution model. Geoffrion and Graves (1974) proposed a multicommodity supply chain design model to optimize product flows from plants to RDCs, RDCs to CDCs, and CDCs to final customers. Work by Brown, Graves, and Honczarenko (1987), Cohen and Lee (1988), and Arntzen and colleagues (1995) models the location and allocation problem as a mixed-integer linear programming problem and provides an efficient heuristic algorithm to solve large-scale problems.

Geoffrion and Powers (1995) examine the evolution of the strategic distribution system design since 1970. Reviews of distribution models with emphasis on supply chain models can also be found in the work of Vidal and Goetschalckx (1997) and Beamon (1998). In a review, Erenguc, Simpson, and Vakharia (1999) emphasize the importance of operational issues, such as lead times in making location/allocation decisions. Melkote and Daskin's (2001) model involves both fixed and arc/variable costs and focuses on the number, location, capacity, and size of warehouses to be set up to maximize profits. However, in this model, manufacturing facilities are not taken into account. Eskigun and colleagues (2005) study the distribution management issue faced by a large-scale automotive firm. Finally, Saourirajan, Ozsen, and Uzsoy (2007) focus on stochastic issues and incorporate lead time and safety stock into their model.

The problem we encountered at GSK required addressing four main questions: (1) How many distribution centers should be opened? (2) Where should the distribution centers be located? (3) What should the capacity of each distribution center be? and (4) How should customers be allocated to distribution centers? Table 1 summarizes the literature that is most relevant to our problem. Our ability to address all four questions in a unified framework provides an important contribution to the literature.

TABLE 1
Distribution Management Literature Most Relevant to the Studied Problem

Method	Questions Addressed	Technique	Strength	Weakness
Amiri (2006)	1, 3, 4	Lagrangian heuristic	Include plants and warehouse decision	Single product type
Brown, Graves, and Honczarenko (1987)	1, 4	Goal decomposition	Where and how much to produce, where to ship from	Focus on solution time
Eskigun et al. (2005)	1, 3, 4	Lagrangian heuristic	Consider lead time, capacity	Single product type: vehicle
Jayaraman and Ross (2003)	1, 3, 4	Mixed-integer programming, simulated annealing	Include cross-docking	Need to know potential cross-docks and warehouses
Melkote and Daskin (2001)	1	Mixed-integer programming	Consider capacity	Demand travels to facilities
Moon and Chaudhry (1984)	2	Integer programming	Introduce various distance-constrained problem	Focus on location only
Swersey and Thakur (1995)	1, 3	Integer programming, set covering problem	Identify location	Single stage, no distribution decision involved

Notes: The numbers in Column 2 correspond to the following questions: (1) how many distribution centers, (2) where to locate, (3) what capacity, and (4) how to allocate customers.

The Continuous Location Problem

The main distinction between our model and the traditional supply chain design is that ours is a "continuous model" rather than a traditional discrete location model. Furthermore, we do not begin with a preset network design. Our continuous model assumes that facilities (e.g., RDCs) can be represented by any point in the Euclidean plane, and travel distances in the mathematical model are calculated by either the Euclidean metric or the Manhattan metric. Conversely, the traditional discrete models, which form the bulk of prior research, assume that facilities can be located only at specific and limited numbers of potential sites. Because of its relevance to our research, we review the continuous model next.

The core of the continuous location problem rests on the Weber problem (Wesolowsky 1993). It determines the coordinates of a single facility, such that the sum of the (weighted) distances $w_l \times d_l(x, y)$ from the facility to the customer at (a_l, b_l) is minimized—that is, $\text{Min}\Sigma_l w_l \times d_l(x, y)$. Among the many measures proposed to determine the proximity between two points on a plane, the Euclidean distance is the simplest and easiest to implement (Anderberg 1973; Gower 1985). For a facility at (x, y) and customer l at (a_l, b_l), the Euclidean distance is computed as $d_l(x, y) = \sqrt{(x - a_l)^2 + (y - b_l)^2}$. The coordinates of the city in which a customer resides can be uniquely identified by the zip code, which matches a specific city.

An extension of the problem that allows for multiple facilities and allocates demands to facilities is the multi-source Weber problem (MWP). Locating multiple facilities simultaneously in a plane to minimize the total transportation cost and to satisfy the demand for many users is a nondeterministic polynomial-time hard problem (Klose and Drexl 2005), and it can be modeled as a nonlinear mixed-integer program as follows:

$$\text{Min}\sum_{l=1}^{L}\sum_{k=1}^{k} w_l d_l(x_k, y_k) z_{kl},$$

subject to $\Sigma_{k=1}^{K} z_{kl} = 1$ for each customer l, where $z_{kl} = \{0, 1\}$ and z_{kl} equals 1 if customer l is assigned to facility k, and x, y are continuous variables.

The main difficulty in solving the MWP arises because the objective function is not convex (Cooper 1967) and can have a large number of local minima. Heuristics are needed to solve large problems and to provide good initial solutions for exact algorithms. In the MWP, it is assumed that the number of new facilities to be located (k) is given. In practice, however, determining the number of facilities is one of the main questions that needs to be answered. Rosing (1992) and Du Merle and colleagues (1999) reformulate the preceding MWP model as a set partitioning problem and use the column generation approach to solve the linear programming relaxed version of the problem. Other variants and extensions of the MWP can be found in Klamroth (2001). To date, all continuous location problems discussed in the literature have been single echelon (i.e., they focus only on one level of supply and demand). However, multi-echelon supply chains (i.e., various levels in distribution network, including suppliers, manufacturers, distributors, retailers, and customers) are needed to carry out the goals of an organization such as GSK. Thus, we develop a model for multiechelon supply chain network design.

Problem Formulation and Methodology

A supply chain network should satisfy customers' demands simultaneously at a desired service level and at the lowest possible cost. To do this, we propose a generic modeling framework that is flexible and general enough to incorporate various constraints such that important location and allocation conditions are taken into account. We call our approach the "continuous supply chain design" (CSCD) problem. This approach results in a realistic nonlinear mixed-integer programming model.

The CSCD Model

The following CSCD model is an extension of the MWP to the multiechelon setting. It focuses on the decisions of location and allocation of RDCs. We describe the notation in the Appendix and formulate the CSCD model as follows:

$$\text{Min } g \times \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{i=1}^{I} d_j(x_k, y_k) \times s_{ijk} + h \times \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{i=1}^{I} d_l(x_k, y_k) \times t_{ikl} \times u_{kl} + \sum_{k=1}^{k} f(w_k) \times z_k, \tag{1}$$

subject to

$$\sum_{k} u_{kl} = 1 \qquad \forall l; \tag{2}$$

$$\sum_{j} s_{ijk} = \sum_{l} t_{ikl} u_{kl} \qquad \forall i,\ k; \tag{3}$$

$$\sum_{k} z_k \le K; \tag{4}$$

$$u_{kl} \le z_k \qquad \forall l; \tag{5}$$

$$\sum_{k} t_{ikl} = D_{il} \qquad \forall i,\ l; \tag{6}$$

$$\frac{\sum_{i}\sum_{l} t_{ikl}}{\sum_{i}\sum_{k}\sum_{l} t_{ikl}} = w_k \qquad \forall k; \tag{7}$$

$$u_{kl},\ z_k = \{0, 1\} \qquad \forall k,\ l; \tag{8}$$

$$s_{ijk},\ t_{ikl} \ge 0 \qquad \forall j,\ k,\ l; \tag{9}$$

$$x_k,\ y_k \text{ are continuous variables.} \tag{10}$$

The CSCD is a large-scale nonlinear mixed-integer programming model. The objective function, Equation 1, minimizes the cost of transportation between plants and RDCs, the cost of shipping between RDCs and CDCs, and the costs of opening and operating RDCs. Constraints in Equation 2 are single-sourcing constraints that restrict a customer's demand for any commodity to be served by a single RDC. Constraints in Equation 3 ensure that all products shipped to an RDC will be shipped to CDCs. Constraints in Equation 4 specify the maximum number of RDCs to open. Constraints in Equation 5 allow CDCs to be assigned to the opened RDCs only. Constraints in Equation 6 ensure that all customer demands are satisfied. Constraints in Equation 7 determine the relative size of warehouse k. Constraints in Equation 8 ensure that u_{kl} and z_k are binary. Variables s_{ijk} and t_{ikl} are nonnegative, as required in the constraints in Equation 9. The coordinates (x_k, y_k) take any positive or negative number without restriction, as the constraints in Equation 10 show.

The CSCD model simultaneously identifies the appropriate sites for RDCs, allocates each CDC to a specific RDC, determines the ideal number of RDCs, determines the size of each RDC, and minimizes the total distribution network costs. In addition to the multiechelon continuous nature, the proposed CSCD model differs considerably from the traditional location and allocation models in several ways. First, information about warehouse capacity (size) is not required, which is different from previous models in the literature (e.g., Amiri 2006; Brown, Graves, and Honczarenko 1987; Eskigun et al. 2005; Jayaraman and Ross 2003; Melkote and Daskin 2001). Note that the size of an RDC is proportional to the total demand assigned to that specific site, and the costs of the RDCs are estimated on the basis of size. In other words, both the size and the location of any RDC depend on the solution of the model and are not provided as parameters. Managerially, this implies that the leasing and operating costs depend on the size of the RDCs opened. Second, no potential RDC locations need to be identified before solving the model, which is different from most of the location selection models in the literature (e.g., Moon and Chaudhry 1984; Swersey and Thakur 1995). Third, prior models have used constant shipping costs to compute the total transportation costs (e.g., Amiri 2006; Elhedhli and Goffin 2005; Erlenkotter 1978; Shen 2005). In this study, transportation cost (parameters g and h in the CSCD model) is a function of the diesel price, distance, weight, line-haul costs, and fuel surcharge, as we discuss in the subsequent subsections. Incorporating these factors makes the model more realistic because these are issues that management must address on a day-to-day basis.

The complexity involved in making the problem formulation more realistic implies that the problem also becomes more difficult to solve optimally. To the best of our knowledge, the CSCD model in this study is the first to integrate a continuous location problem and a comprehensive multiechelon supply chain network design problem, both of which are classified as nondeterministic polynomial-time hard problems. Problems in this category cannot be optimally solved in a reasonable amount of time (polynomial time), often taking days, weeks, and more, depending on the scale of the problem. As a consequence, exact optimal solution methods are restricted to small-scale problems, and such problems often end up being unrealistic. Thus, effective heuristic algorithms are developed to provide near-optimal solutions within minutes. Next, we describe how we solve the CSCD problem using a heuristic.

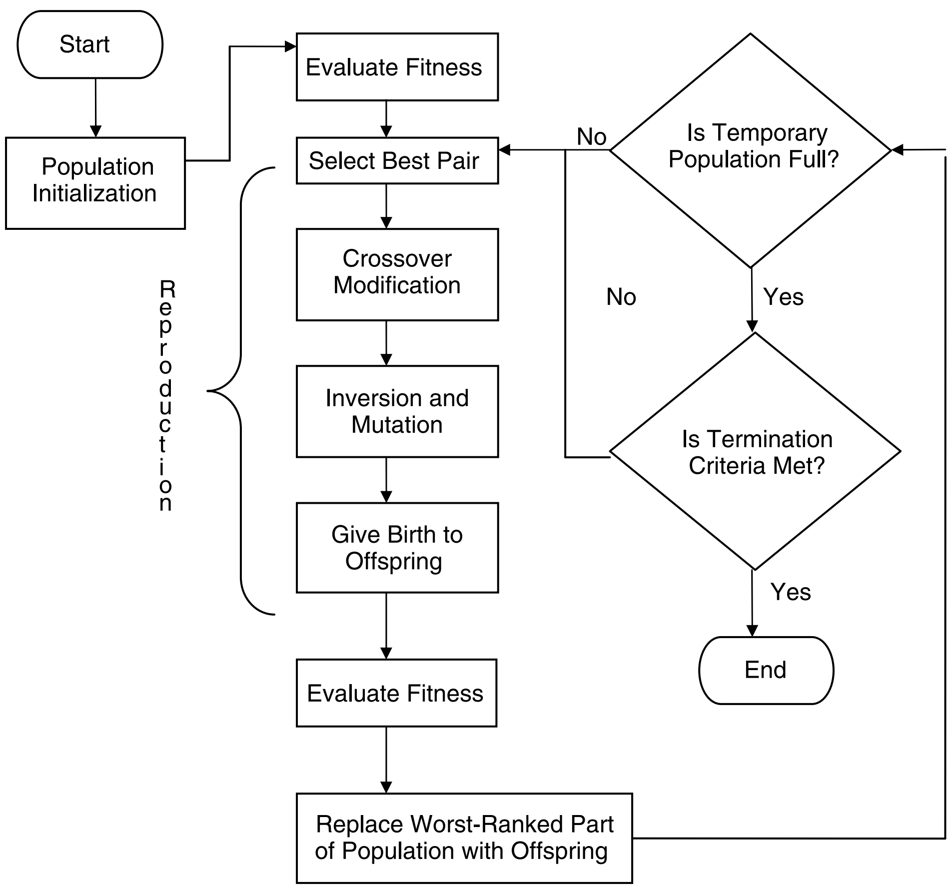
Solution Approach

Many heuristics of different accuracy and speed have been suggested to solve the location and allocation problem. Most problems are solved by modern heuristic procedures, which do not guarantee an optimal solution. Our goal is not to pursue the exact mathematical optimum but rather to solve the problem efficiently and realistically. In other words, we want an algorithm that can generate a relatively good solution within a reasonable amount of time (i.e., minutes or hours, not months or years). The genetic algorithm in Evolutionary Solver can serve such a purpose. The Evolutionary Solver (Ashlock 2005) is a Microsoft Excel add-in tool found in Premium Solver (Frontline Systems 2007).

The genetic algorithm approach, introduced by Holland (1975), is a global search heuristic procedure that incorporates processes inspired by evolution ideas in biology, such as initiation, selection, reproduction (crossover and mutation), and termination (see Figure 1). A series of steps are used to solve the CSCD model. First, the genetic algorithm starts by randomly generating a large set of candidates to form initial solutions, called a "population," and then evaluates the fitness (quality) of each individual candidate (i.e., solution) in the population. Whether an individual solution is fit depends on whether it can satisfy all constraints and the solution quality. Second, the genetic algorithm selects two best-ranking individual solutions to reproduce through crossover and mutation (genetic operations) and to give birth to offspring (i.e., pairs of the existing population to create offspring for the next generation). Third, the genetic algorithm evaluates the fitness of each offspring, replacing the worst-ranked part of population with the best offspring. Therefore, the population evolves and becomes more fit. Fourth, the genetic algorithm occasionally makes a random change by substituting a member in the population by a random value. Such a mutation can create offspring that are far removed from the rest of the population to avoid being stuck at a local optimum. Fifth, the genetic algorithm continues creating new generations until no improvements occur in several successive generations. The algorithm terminates, and the best result found becomes the solution.

In applying the genetic algorithm in Evolutionary Solver, we use large values in maximum subproblems and maximum feasible solutions to extend the search. Increasing the population size and/or mutation rate helps improve with searches that were trapped in local optimum. Con-

FIGURE 1
The Flowchart of Genetic Algorithm



versely, decreasing the tolerance and/or increasing the maximum time without improvement allows for a longer search and a chance to improve results. All these are critical user inputs to the software program to help detect a better solution.

Although the genetic algorithm does not guarantee an optimal solution, Wen and Iwamura (2008) find that it is likely for the genetic algorithm to end up with a solution that is close to optimal. To improve the chance of obtaining the best solution, a standard optimization model can be run, starting with the final solution obtained by the Evolutionary Solver, to seek an opportunity for enhancing the solution quality.

Applying the CSCD Model to GSK

A real-life case study at GSK provides the setting for implementing our model. The CSCD model has successfully handled the problem, and the company's marketing and logistics department recognizes the validity of the results. Because of confidentiality concerns, some numerical results reported here are disguised.

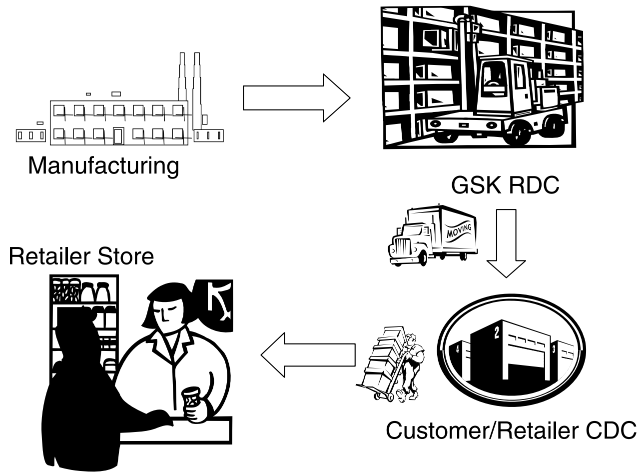
The Distribution Network of GSK

GlaxoSmithKline's supply chain network consists of suppliers, manufacturers, warehouses, and retail distribution channels, which are synchronized to acquire raw materials, produce finished goods, and distribute products to warehouses and retailers. Currently, most products are produced by four manufacturing plants in the United States and more than 30 contractors around the world. They manufacture 60 brands and more than 1200 different items, representing approximately 4 billion packs per year. Together with two copacking facilities, they form the supplier network of the company's consumer health care products. The firm has four RDCs in the United States to ship its products to more than 400 retail accounts, each of which has multiple CDCs. In all, there are 25,000 "ship-to" locations (i.e., retail stores). Annually, more than 80,000 customer orders are handled, and 20 million cases of products are shipped. Figure 2 provides the direction of the inventory flow. The locations of current manufacturing plants and RDCs appear in Figure 3.

As Figure 3 shows, GSK currently has four RDCs located in Fountain Inn, S.C.; Memphis, Tenn.; Hanover, Penn.; and Fresno, Calif. These RDCs receive inbound shipments from manufacturing facilities located in Aiken, S.C.; Clifton, N.J.; Memphis, Tenn.; and St. Louis; Mo. No products are stored in the manufacturing facilities. If an RDC is short of a product, another RDC is allowed to share inventory and ship to the RDC in need. However, management has deemed that such an activity wastes resources and thus is strongly discouraged. Therefore, interwarehouse shipping is avoided and should not be considered in our model.

After an RDC receives products, it ships them to the corresponding CDCs. When the CDC receives the products, the customer takes possession and title of the goods. These CDCs are spread around the United States, with more CDCs on the East Coast than the West Coast. The existing

FIGURE 2
Distribution Network of GSK



distribution network of GSK has been employed for more than a decade and is a result of incremental changes as the company has grown. A major opportunity for improvement of the distribution network arose when the current leases on some RDCs were due for expiration and renewal. The firm had the option either to renew some of the leases or to look for alternative sites if necessary. In addition, management was concerned with the rising costs of fuel, a key cost factor. During the second half of 2005, diesel prices rose 46.5%, from an average of \$2.15 to \$3.15 per gallon, and they reached \$3.80 in March 2007. The executives mandated that any distribution network design include the costs of energy as a key decision factor. In other words, the firm wanted to relocate its RDCs such that a potential change in the environment, such as customer demand and diesel prices, would have the smallest possible impact on the distribution costs.

Determining the Number, Sites, and Capacities of the RDCs

We obtain the model solution with the Evolutionary Solver, which determines the geographical coordinates of RDCs, which are then converted into zip codes and used to identify the corresponding cities. Specifically, the CSCD model determined that the optimal warehouse (RDC) coordinates were (33.5, –96.9), (37.1, –118.6), (33.4, –83.8), (41.6, –87.9), and (40.2, –76.3). These coordinates corresponded to zip codes 76233, 93513, 31085, 60441, and 17545, representing five cities: Collinsville, Tex. (RDC 1); Big Pine, Calif. (RDC 2); Shady Dale, Ga. (RDC 3); Lockport, Ill. (RDC 4); and Manheim, Penn. (RDC 5). The model also assigned each CDC to one of these RDC locations to minimize the total costs. Many of GSK's retailers have multiple CDCs that are geographically dispersed throughout the United States; therefore, different CDCs for the same retailer (e.g., Wal-Mart) can be served by different RDCs.

We summarize the solution to GSK's RDC location and allocation problem in Table 2 and map this onto U.S. cities in Figure 4. In Table 2, the percentages show the relative

FIGURE 3
RDCs and Manufacturing Plants for GSK

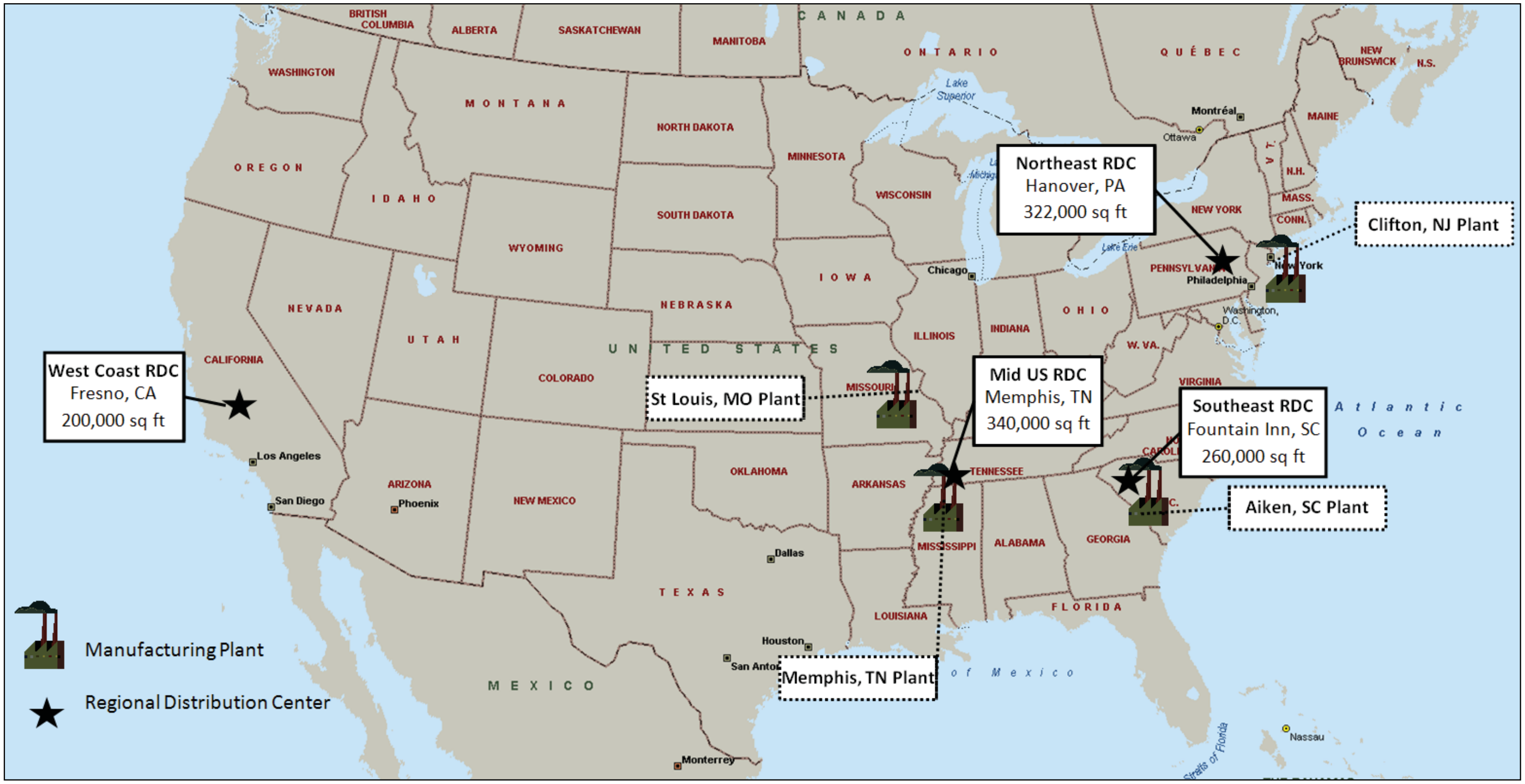


TABLE 2
The Results of Applying the CSCD Model to the GSK Data

Chosen City, State	Zip Code	Relative Size of Warehouse Capacity
Collinsville, Tex.	76233	21%
Big Pine, Calif.	93513	21%
Shady Dale, Ga.	31085	24%
Lockport, Ill.	60441	19%
Manheim, Penn.	17545	16%

size of each warehouse. Each value is derived from Equation 7. Figure 4, Panel A, shows the CDCs that are to be served by each proposed RDC on a U.S. map. Circles of the same shade are CDC locations that are assigned to the same RDC. Figure 4, Panel B, shows the old and new RDCs.

Deriving the Unit Transportation Costs on the Basis of the Cost Structure of Freight

To understand the cost structure of freight for this firm, it is important to distinguish between full-truckload (TL) and less-than-truckload (LTL) shipping. By TL, we refer to shipments of a full truckload to obtain economies of scale, and by LTL we refer to the shipments with a relatively small freight that does not fill the truck. Truck drivers are expected to transport freight at an average rate of 47 miles per hour (this factors in traffic jams and queues at intersections) all the way to destination. The advantage of TL carriers is that the freight is never handled en route, giving a more predictable delivery time, whereas an LTL shipment may be transported on several different trailers to achieve higher volume efficiency. The main advantage of LTL shipment is that it is much cheaper than TL. Many, if not all, carriers view themselves as either primarily LTL or primarily TL carriers. GlaxoSmithKline's manufacturers use TL to ship from plants to RDCs to ensure efficiency and use LTL to ship from RDCs to CDCs to customize clients' needs.

Thus far, researchers (e.g., Amiri 2006; Elhedhli and Goffin 2005; Shen 2005) working on location and allocation problems have used average unit shipping costs to compute the total transportation costs. However, in reality, freight costs differ significantly between TL and LTL shipments. Equations 11–15 distinguish the TL and LTL costs and identify the underlying transportation cost structure for GSK.

When the quantities are shipped as containerized in truckloads or when the transportation cost structure charges a truckload minimum for partial quantities, product distribution cost mainly becomes a function of the distance traveled. However, under a partial-load price structure, it is common to express the pricing in terms of load distances, such as ton–mile, where a ton–mile is the amount of transportation activity to move a ton of material over a distance of one mile. To keep our model general, we express the transportation activity in load distances because the pure distance-based approach is a special case of the general model when the minimum charged quantities are in truckloads.

For both TL and LTL types of transport, fuel surcharge and line-haul charges are the two main costs. The total freight cost can be formulated as follows:

$$\text{(11)} \quad \text{Total transportation cost} = \text{Demurrage} + \text{Line-haul cost} + \text{Fuel surcharge}.$$

All the information needed for Equation 11 is obtained by GSK in advance. Plugging these data into Equation 11 gives the LTL and TL freight costs, which correspond to g and h in the CSCD model. The demurrage in Equation 11 is the cost imposed as compensation for the detention of a carrier taking longer than the normal time needed to load and unload a truck. Line-haul costs are basic charges for long-distance moves, which are usually calculated on the basis of mileage and weight of the shipment. Fuel surcharge is an additional per-shipment fee that carriers impose when fuel prices are above predefined levels.

Using the historical shipment data provided by GSK and the characteristics of customer orders, we express line-haul costs as regression models (see Equations 12 and 13). For the LTL regression equation, the independent variables, distance ($p < .0001$) and weight ($p < .0001$), were both statistically significant. The model was adequate with a high R-square value (98.2%). Conversely, for the TL regression equation, only distance was significant ($p < .0001$), and the model was able to account for a high portion of variation ($R^2 = 91.3\%$).

$$\text{(12)} \quad \text{LTL: Line-haul} = \$12.79 + .06\,(\text{Distance}) + .056\,(\text{Weight}), \text{ and}$$

$$\text{(13)} \quad \text{TL: Line-haul} = \$399.48 + 1.0268\,(\text{Distance}).$$

In these equations, we do not include additional variables, such as size, type, and volume of customer order, because we found them to be insignificant in determining the line-haul costs. In both equations, distance is statistically significant, but weight is significant only in the LTL situation. We expect weight to be insignificant in the TL equation because the truck is always full, and in general, fully loaded trucks weigh similarly in this case.

In addition to line-haul costs, carriers also impose a fuel surcharge cost. Using the Department of Energy's Diesel Fuel Index, we estimate the fuel surcharge costs for LTL and TL in Equations 14 and 15:

$$\text{(14)} \quad \text{LTL: Fuel surcharge} = (\{[.3(\text{Diesel price} - 1.15)/.05] + .7\}/100) \times \text{Line-haul cost, and}$$

$$\text{(15)} \quad \text{TL: Fuel surcharge} = (.2 \times \text{Diesel price} - .2298) \times \text{Miles traveled}.$$

Equation 14 shows that fuel surcharge is the product of the surcharge rate and the line-haul cost for LTL shipment. The surcharge rate in Equation 14 ({[.3(Diesel price – 1.15)/.05] + .7}/100) reflects that for each $.05 increase over $1.15 in the diesel price, the fuel surcharge will increase by .3% of the line-haul cost. In the calculation, this is rounded down for convenience. For TL, the diesel price and miles

FIGURE 4
Number and Sites of RDCs and Allocation of CDCs to RDCs

A: The Five RDC Locations and the Allocation of CDCs to RDCs[a]

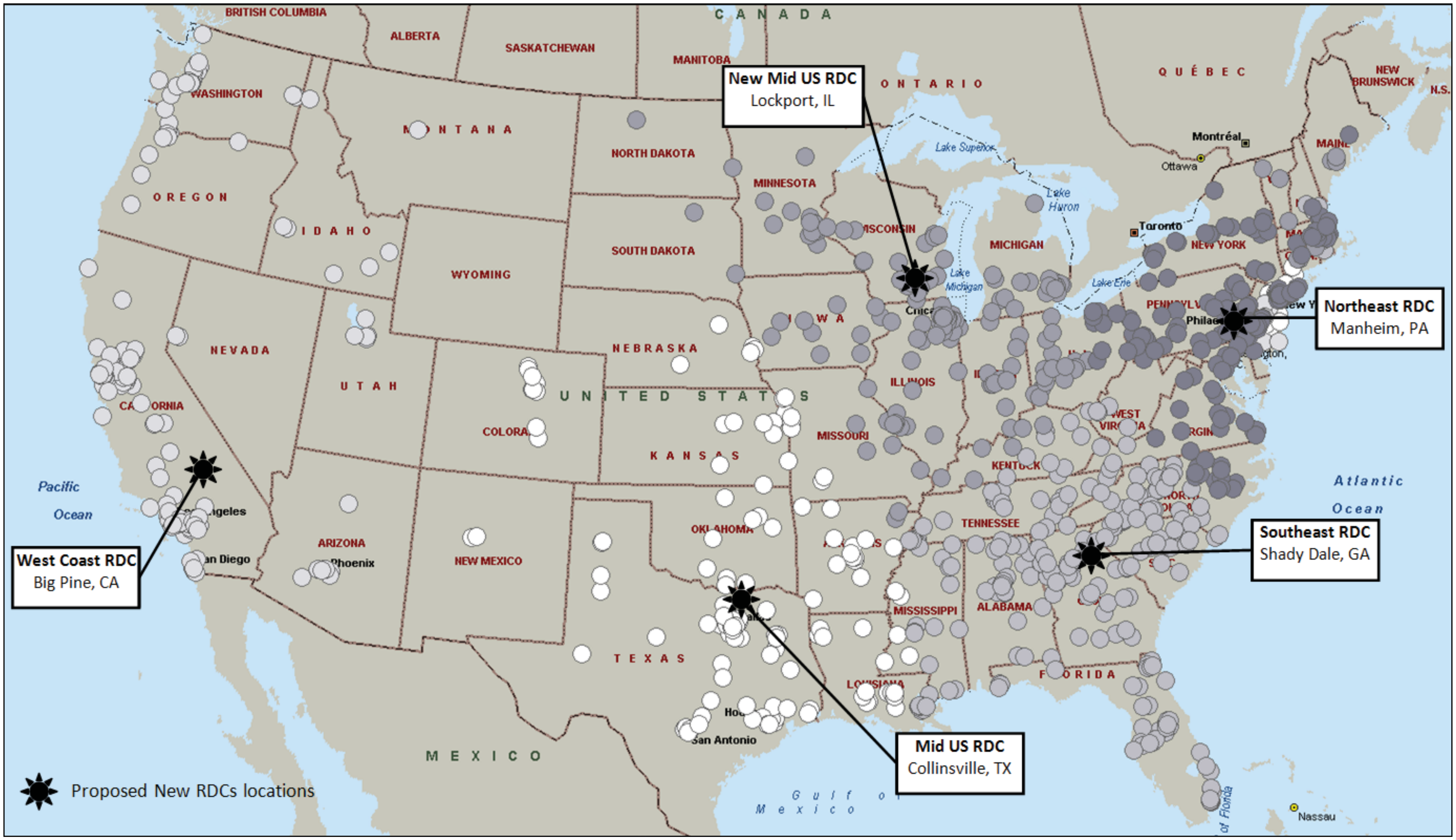
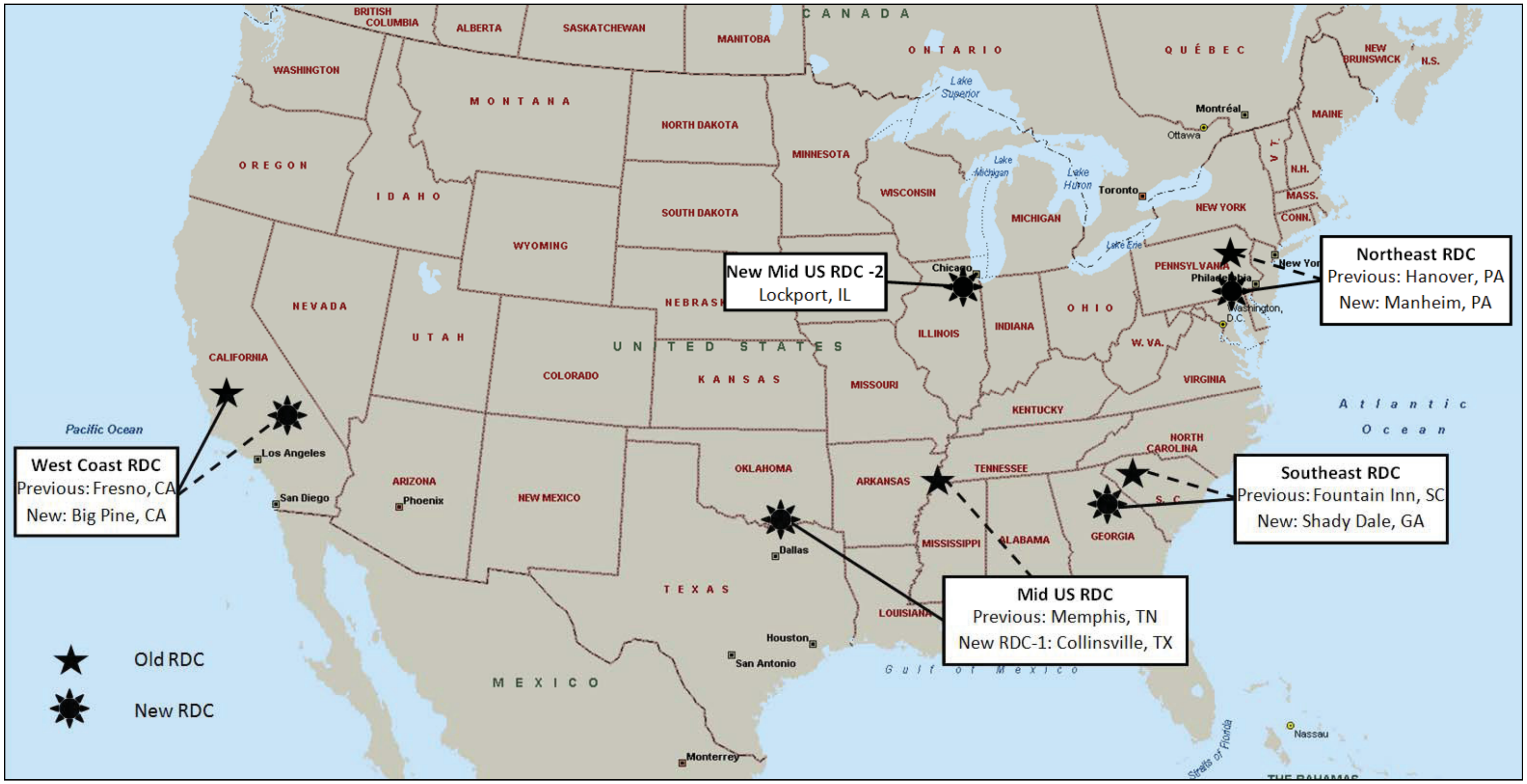


FIGURE 4
Continued

B: Previous RDC Locations and the New Five-RDC Solution Locations[b]



[a]Small circular dots represent existing CDC locations in the United States, and stars with city names and state represent the proposed new five RDC locations.
[b]Relative capacities of the plants are as follows: Collinsville, Tex. (21%); Big Pine, Calif. (21%); Shady Dale, Ga. (24%); Lockport, Ill. (19%); and Manheim, Penn. (16%).

traveled define the fuel surcharge; again, weight is not an important factor.

The parameters estimated in Equations 11–15 help management better understand the relationship among the components of transportation cost. GlaxoSmithKline can plug in these components and quickly and approximately estimate the transportation costs for each of the items it ships annually.

The cost function of RDCs is a function of the size of the warehouse: $676 \times (100 \times \text{Size})^{1/2}$, which suggests that there are economies of scale. If we combine the transportation costs of Equation 11 with warehouse cost function, the total distribution network costs for CSCD can be derived.

Total Distribution Network Costs and Service Level

Although having fewer RDCs lowers the fixed and operating costs associated with the warehouse management, it increases both the inbound and the outbound transportation costs as a result of longer delivery distance. Because they are likely to be inversely related, it is important to balance warehouse costs and transportation costs. We use the firm's shipping data, fuel price prevailing during the data period, warehouse leasing and operating costs, and TL and LTL cost equations to compute the total costs of the distribution network. The percentages of customers who can be served within 100-, 250-, 500-, and 1000-miles radius are also computed.

Table 3, Panel A, summarizes the annual distribution costs for the firm. For the current four-RDC system, the TL transportation cost is $12.27 million (9.14 + 1.87 + 1.26), and that of the LTL is $5.98 million (5.44 + .41 + .13). Together, the transportation cost is $18.25 million. When the warehouse costs are taken into consideration ($15.41 million), the current annual distribution network cost is $33.66 million.

The unit transportation costs, g and h, in Equation 1 are $1.36 and $1.70, respectively. We derive parameters g and h as follows: We obtain the value for g by dividing the total TL shipping cost by the weighted miles traveled. Similarly, dividing the total LTL shipping cost by the total LTL weighted miles traveled gives the cost h in the CSCD model. As we discussed previously, GSK's actual transportation cost is $18.25 million, which includes $5.98 million for the TL shipping and $12.27 million for LTL shipping. Dividing the $5.98 million TL shipping cost by the TL weighted miles of 4,399,253 gives a value of $1.36 for g. Similarly, we obtain h by dividing the $12.27 LTL cost by 7, 217,000 mile–ton traveled for LTL shipping.

Because both g and h are linearly related to weighted travel distance in Equation 1, they imply that traveling one mile–ton distance by TL will increase the overall transportation costs by $1.36. Conversely, a mile–ton of LTL travel will increase the LTL transportation costs by $1.70. Recall that TL is used to go from plants to RDCs, and LTL is used to go from RDCs to retailers' CDCs.

Table 3, Panel A, shows that the total transportation costs vary with the total number of warehouses and that there is a trade-off between warehouse expenditure and transport spending. For example, with only two RDCs, the transportation costs are higher ($22.25 million), though the warehouse costs are lower ($12.32 million). The five-RDC distribution network represents the optimal balance because the total cost decreases from the two-RCD solution and reaches a minimum at five RDCs, beyond which it increases again. Under the five-RDC system, the transportation cost is estimated to be $14.99 million annually, a savings of $3.26 million (17.9%) from the current four-RDC network system. However, by adding one more RDC, an incremental warehouse cost of $1.27 million ($16.68 – $15.41 million) will be incurred. Combining the transportation and warehouse costs, we obtain a total cost of $31.67 million with the five-RDC solution, rather than the current cost of $33.66 million. Thus, this new network design saves a total of $1.99 million (or 6%) in distribution costs per year. In addition, as we explain next, there are also savings and benefits emanating from improved delivery time.

Improving Customer Service by Shortening Delivery Radius and Time

The percentage of orders that can be delivered to the required CDC location within one day is an important metric of customer responsiveness and customer service for the firm. However, according to transportation law, truckers in the United States can only drive a maximum of 11 hours (11 hours $\times$ 47 miles/hour = 517 miles/day $\approx$ 500 miles) after 10 consecutive hours off duty. When the driving distance exceeds 500 miles, it takes more than one day to deliver the products, which is undesirable. As such, a desirable distribution network design would be one in which the majority of the CDCs are within 500 miles of their serving RDCs.

Table 3, Panel B, shows that with the five-RDC plan, the company can ship 86.2% of customer orders within one day.[1] Compared with the current service level of 61.41% one-day deliveries, it represents a 40.4% improvement. This improvement is primarily what makes the five-RDC design attractive to GSK. Furthermore, with the proposed five-RDC option, only 2.2% of the orders will take longer than two days (more than 1000 miles distance) to deliver. For small orders that take longer than one day, the company may opt to expedite through one-day air service if needed.

In summary, the proposed five-RDC distribution network system not only decreases the total cost by 6% but also improves the one-day delivery rate by 40.4%. As we expected, management at GSK was keen to adopt the proposed system. However, we still need to investigate whether the proposed five-RDC solution remains optimal under different environments (e.g., changes in fuel price, demand quantity, RDC costs, number and locations of customers, and required service level). We answer these questions by conducting sensitivity analysis.

[1]Within-one-day deliveries correspond to deliveries within a 500-mile radius.

TABLE 3
Total Distribution Costs and the Service Level

A: The Total Distribution Costs for K = 2, 3, …,10 Warehouses[a]

	LTL				TL						
	Line-Haul ($)	Fuel ($)	Demurrage ($)	Total Cost for LTL ($)	Line-Haul ($)	Fuel ($)	Demurrage ($)	Total Cost for TL ($)	Transportation Cost ($)	Warehouse Fixed and Operating Costs ($)	Total Costs ($)
Current 4-RDC system	9.14	1.87	1.26	12.27	5.44	.41	.13	5.98	18.25	15.41	33.66
2 RDC	11.23	2.34	1.26	14.83	6.81	.48	.13	7.42	22.25	12.32	34.57
3 RDC	10.07	2.17	1.26	13.5	5.75	.45	.13	6.33	19.83	13.99	33.82
4 RDC	8.94	1.86	1.26	12.06	4.74	.41	.13	5.28	17.34	15.41	32.75
5 RDC	**7.9**	**1.68**	**1.26**	**10.84**	**3.68**	**.34**	**.13**	**4.15**	**14.99**	**16.68**	**31.67**
6 RDC	6.99	1.52	1.26	9.77	3.31	.31	.13	3.75	13.52	18.67	32.19
7 RDC	6.78	1.36	1.26	9.4	2.87	.27	.13	3.27	12.67	19.95	32.62
8 RDC	6.63	1.34	1.26	9.23	2.65	.23	.13	3.01	12.24	21.26	33.5
9 RDC	6.43	1.34	1.26	9.03	2.28	.2	.13	2.61	11.64	22.33	33.97
10 RDC	6.2	1.34	1.26	8.8	2.01	.19	.13	2.33	11.13	23.38	34.51

B: The Percentage of Customer Orders That Are Served by Warehouses Located Within Specific Radius[b]

	Percentage of Customer Orders Served for Each CDC–RDC Distance Range Under Each Scenario (%)						
	Less Than 100 Miles	100–250 Miles	250–500 Miles	500–1000 Miles	More Than 1000 Miles	Total Percentage of Customers Served in One Day (%)	Total Percentage of Customers Served in More Than One Day (%)
Existing 4-RDC system	3.18	23.26	34.97	33.68	4.91	61.41	38.59
2 RDC	.10	6.70	41.00	47.30	4.90	47.80	52.20
3 RDC	.60	14.60	42.10	37.90	5.00	57.30	42.70
4 RDC	3.40	18.10	42.50	31.70	4.30	64.00	36.00
5 RDC	**8.80**	**27.70**	**49.70**	**11.60**	**2.20**	**86.20**	**13.80**
6 RDC	10.30	37.30	41.10	9.40	1.90	88.70	11.30
7 RDC	13.50	40.20	38.20	6.90	1.20	91.90	8.10
8 RDC	17.70	39.10	31.40	5.10	.80	88.20	11.80
9 RDC	21.30	40.30	35.40	2.90	.10	97.00	3.00
10 RDC	24.20	43.40	32.00	.40	.00	99.60	.40

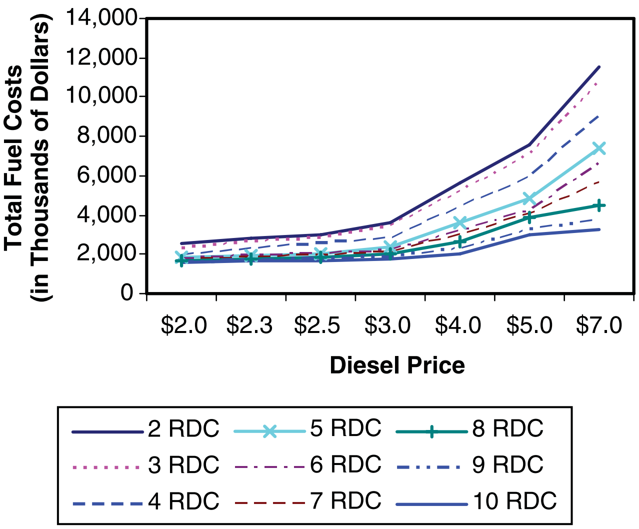
[a]All dollar figures are in millions. We calculated all figures for 12 months. We assumed diesel price to be $2.3. The best solution (the 5-RDC scenario) appears in bold.
[b]Customers that are within 500 miles radius can be served in one day. Diesel price per gallon is assumed to be $2.3. The best solution (the 5-RDC scenario) is given in bold.

Sensitivity Analysis and Heuristics Comparison

A major obstacle in the design of a supply chain network is the uncertainty underlying the supply chain parameters. The stochastic nature of distribution networks makes most traditional analytical models either overly simplistic or unsolvable. Solvable models may not be robust and may become invalid under different business environments. The sensitivity analyses we perform examine the capability and robustness of the proposed model in handling variability.

Sensitivity analysis examines how the results of a model vary with changes in model inputs (Meepetchdee and Shah 2007). A model is said to be sensitive to an input if changing an input variable changes the model output (i.e., the optimal solution).

FIGURE 5
The Impact of Changes in Diesel Price on the Total Fuel Costs Under Various Scenarios



Changes in Fuel Price

To understand the impact of fuel price uncertainty on the efficiency of the distribution network, we conducted a sensitivity analysis with respect to various levels of fuel prices. Figure 5 shows the results of such an analysis. Because diesel price directly influences the fuel surcharge, we computed the combined fuel surcharges for both LTL and TL in the sensitivity analysis.

Table 4 and Figure 5 show that the total fuel surcharge is inversely proportional to the total number of RDCs. When fuel price is low, the difference in total fuel surcharge costs does not differ significantly; however, the difference increases with a rise in fuel price. To mitigate the negative impact of increasing fuel price, it is desirable to adopt more warehouses. However, recall that warehouse opening and operating expenses are high, and they constitute a significant portion of the overall distribution network costs. Thus, the savings in fuel costs from having more RDCs may be offset by additional warehouse costs. A desirable decision hinges on balancing the transportation costs and the warehouse expenses. When both costs are taken into consideration, the CSCD model recommends the five-RDC solution when fuel prices fall in the range of ($2.3, $5), a highly likely event. When fuel cost is less than $2.3, the four-RDC solution reaches the lowest total cost. If the fuel cost is within the range of ($5, $6), the six-RDC solution is the choice. A price of $6 or above makes the seven-RDC solution the best option.

Changes in Demand Level

To examine the impact of demand changes in CDCs on model performance, we generate random demand by using current demand $\times$ α, where α is a random number uniformly distributed between .5 and 1.5 (i.e., $\alpha \sim U[.5, 1.5]$). When α is close to 1, it indicates a small variation. When $\alpha = .5$, it implies that the demand is only half the original demand quantity. Conversely, when $\alpha = 1.5$, it shows that the demand has increased 50% from the previous amount. We found that for $.80 \leq \alpha \leq 1.15$, a range in which demand is likely to change between a 20% decrease and a 15% increase, which is a likely fluctuation GSK may come

TABLE 4
Total Fuel Surcharge at Different Diesel Price Levels Under Different Scenarios

	Total Diesel Price GSK Incurs ($)						
Fuel Cost ($)	**2.0**	**2.3**	**2.5**	**3**	**4**	**5**	**7**
Existing 4-RDC System	1,958	2,282	2,498	2,907	4,621	5,955	9,182
2 RDC	2,586	2,820	2,976	3,646	5,619	7,593	11,539
3 RDC	2,320	2,620	2,820	3,393	5,228	7,063	10,734
4 RDC	1,922	2,276	2,512	2,844	4,383	5,921	8,999
5 RDC	1,844	1,970	2,054	2,346	3,616	4,885	7,425
6 RDC	1,765	1,904	2,003	2,204	3,213	4,215	6,588
7 RDC	1,702	1,856	1,907	2,089	2,987	4,022	5,672
8 RDC	1,658	1,741	1,875	1,992	2,645	3,854	4,531
9 RDC	1,604	1,698	1,765	1,854	2,330	3,481	3,458
10 RDC	1,587	1,630	1,689	1,794	2,050	2,963	3,216

Notes: All numbers are derived from Equations 14 and 15. The final decision on the number of RDCs to employ must include these numbers and the RDC costs.

across, the five-RDC solution remains the best. When $1.15 < \alpha \leq 1.30$ (demand is increased by 15%–30%), the six-RDC solution gives the lowest cost solution. When $1.3 < \alpha \leq 1.5$ (demand is increased by 30%–50%), the seven-RDC solution is the best. Conversely, we found that when customer demand is reduced by 20%–50% (i.e., $.50 \leq \alpha < .80$), the four-RDC solution minimizes the overall costs.

Changes in Warehouse Costs

To examine whether the five-RDC solution remains valid when warehouse costs change, we varied the RDC costs function $[\beta \times (100 \times \text{Size})^{½}]$ by changing the coefficient β from the original value (676) to values within the range (500, 900) in increments of 50. We found that under a small change (i.e., when $600 \leq \beta \leq 800$), the five-RDC solution is still optimal. When $\beta < 600$, the six-RDC solution is the best solution. When warehouse costs increase significantly (i.e., $\beta > 800$), the four-RDC solution is optimal. Because RDC costs do not typically change drastically over time and often fall within the range of $600 \leq \beta \leq 800$, the five-RDC decision still remains the best option.

Changes in the Number and Location of CDCs

We also examined the impact of changes in the number and location of CDCs. We found that when the number of CDC locations varies within ±9% of the existing CDC locations or when the location coordinates vary within ±13% of the existing CDC locations, the five-RDC decision remains optimal.

The preceding discussion shows that the five-RDC solution recommended by the CSCD model is robust and remains the best choice under realistic and reasonable situations. For general tests of different problems combining multiple factor changes simultaneously, we conduct the experiments reported next.

Changes in Service Level

A practical and valuable question is how many RDCs would be optimal for 100% delivery of customer service within one day given the nonlinear mixed-integer model. Recall that one-day delivery is important given the constraint from regulators that drivers can travel no more than 500 miles in a given day. The problem can be modeled by adding a constraint such that all the CDCs will be within a distance of 500 miles of their served RDCs (i.e., $d_l[x_k, y_k] \leq$ 500 miles). We incorporate this constraint into the original CSCD model and solve the model again using Evolutionary Solver.

The results show that for GSK, ten warehouses would be needed to achieve a 100% service level (i.e., delivery to all customers within one day). In reality, however, GSK is not interested in a 100% delivery within one day because of excessive costs. As with most firms, GSK would use air freight to expedite some shipments, even though expediting increases the transportation costs. This is still much more economical than opening extra warehouses. Management at GSK regarded opening another five RDCs to achieve an improvement of 13.8% in one-day delivery as not justifiable. However, in other situations (e.g., emergency medical supplies) and for other firms, the trade-off between service improvement and extra RDCs may be worthwhile.

Comparing the Performance of Different Heuristics for CSCD Problem

Except for some small problems, there is no guarantee that the theoretically optimal solutions (cost levels) will be obtained in real life because of the stochastic nature of the distribution networks. Managers are often faced with the need to find high-quality solutions to difficult problems, such as CSCD. Although preferred because of their combinatorial nature, larger-scale problems often cannot be solved optimally within a reasonable time, as we mentioned previously. Thus, managers regularly turn to heuristics such as the genetic algorithm to search for solutions. The genetic algorithm–based Evolutionary Solver is a heuristic approach in which an optimal solution is not guaranteed. This is the undesired consequence of most heuristic search approaches, though many researchers have reported that intelligent heuristics find extremely good solutions (Eiben and Smith 2007; Menon 2004).

To examine the effectiveness of the genetic algorithm, we conducted an empirical comparison of the genetic algorithm with two other heuristics. The first is the simulated annealing heuristic, which is a randomized local search method that approximately solves an optimization problem. Simulated annealing navigates the search space by exploring the performance of the neighbors of the current solution. A superior neighbor is always accepted. An inferior neighbor is stochastically accepted on the basis of the difference in quality and a temperature parameter. The temperature parameter is modified as the algorithm progresses to alter the nature of the search. Jayaraman and Ross (2003) study the applicability of simulated annealing and suggest that simulated annealing is an effective and useful solution approach to complex problems involved in supply chain management.

The second is the relocation search method proposed by Brimberg and colleagues (2000). It constructs its neighborhood as the set of points obtained by a given number of facility relocations. Instead of visiting all points in the interchange neighborhood, a strategy referred to as drop-and-adds is used. Brimberg and colleagues use drop-and-adds to determine which facility to drop first and where the best site is to reinsert it next. The steps are as follows: (1) find an initial solution, (2) drop a facility according to the least useful strategy, (3) reinsert the RDC at an unoccupied customer location according to the most useful strategy, and (4) use Cooper's (1964) algorithm and the modified set of RDCs to find a local minimum. If it improves, save the new currently best solution and return to the second step; otherwise, stop.

To compare the performance of the three heuristics, we test assorted problems by varying the parameters of the CSCD. First, we generate the unit fuel price from uniform distribution (i.e., U[\$2.3, \$5]) to provide inputs for deriving the unit transportation cost of g and h in the objective function of CSCD. Note that g is the unit cost of shipping from plants to RDCs, and h is the unit cost of shipping from

RDCs to CDCs. Second, demands are generated from another uniform distribution multiplied by the demand (i.e., U[.8, 1.15] × current demand), and the fixed cost of RDCs is generated from $\beta \times (100 \times \text{Size})^{1/2}$, where $\beta \sim U[600, 800]$. Finally, we vary the number of customers within ±9% of the existing customer number, and the location coordinates vary within ±13% of the existing CDC locations. Overall, 500 problems are generated. For each problem, five random initial solutions are compared. The best of the five results is chosen as the solution for the specific problem. We repeated this procedure for each of the 500 problems under each of the three heuristics.

An example of the five solutions generated for 1 of the 500 problems appears in Figure 6. We observe similar outcomes and patterns in the remaining problems. Therefore, we conclude that costs and one-day 500-mile delivery performance are not significantly different among the three heuristics. No one single algorithm dominates another one. The genetic algorithm, simulated annealing, and relocation search heuristics all perform comparably. However, the genetic algorithm heuristic can be solved in Excel using the Evolutionary Solver add-in. Thus, we consider it easy and straightforward for management implementation.

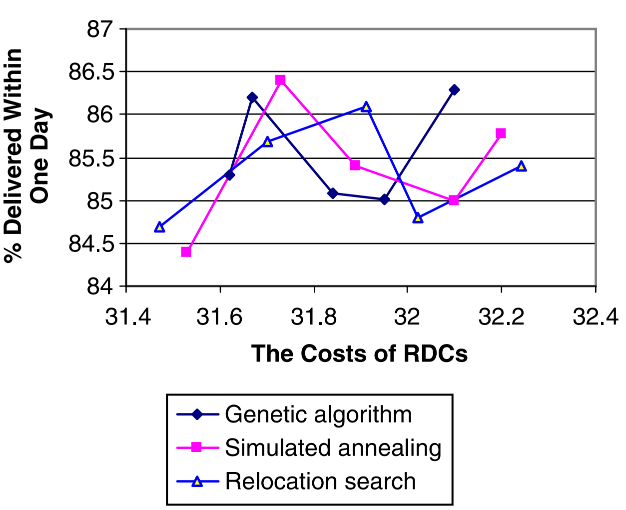
Discussion

Marketing scholars have long understood the importance of location analysis in determining the marketing success of firms. While researchers in marketing have focused on retail locations (Chan, Padmanabhan, and Seetharaman 2007; Ghosh and Craig 1983; Rust and Donthu 1995), location analysis of manufacturing and distribution systems has been assumed to be the domain of operations research. Yet it has become increasingly clear that companies striving to achieve marketing success in their retail operations must incorporate strategic supply chain planning—distribution networks—into their decisions. In addition to mitigating the deleterious impact on customers from outcomes such as stockouts (Anderson, Fitzsimons, and Simester 2006), such an approach can enhance the productivity and profitability of both the retailer and the manufacturer (Rust et al. 2004). To this end, this article has proposed a continuous, uncapacitated, deterministic supply chain network model. An optimal distribution network model, such as the CSCD, can substantially reduce distribution expenditure while enhancing service levels through continuous supply and reduced stockouts. Traditional distribution network models try to minimize total distribution cost on the basis of a few predefined alternative locations. In this study, different from traditional models, we are not given candidate locations; the entire U.S. map provides near-infinite potential warehouse locations. By incorporating key information into the distribution system, including zip codes, fuel price, and TL and LTL freight structures, and by converting mileage to carrier time in transit, we can comprehensively examine and compare all cities in the United States.

Our proposed approach has been implemented by GSK, a major pharmaceutical firm, in conjunction with the marketing department. GlaxoSmithKline views its distribution network as a key element of its marketing strategy. Using a decision calculus approach (Little 2004), GSK can redesign the distribution network to reduce distribution costs while significantly increasing the one-day service level. Locating the five RDCs in the recommended locations (Collinsville, Tex.; Big Pine, Calif.; Shady Dale, Ga.; Lockport, Ill.; and Manheim, Penn.) offers GSK the opportunity to attain the most economical network design, while providing an 86.2% next-day service level to its customers. In addition, 97.8% of all customer orders will be complete and delivered within two days. Currently, GSK is in the process of implementing these changes in its distribution networks.

The managerial implications of this study, beyond GSK, are threefold. First, distribution networks play an important role in simultaneously enhancing effectiveness and efficiency of marketing systems in general. Instead of taking a narrow view of marketing to exclude distribution and manufacturing, a more integrative and comprehensive view is warranted. An optimal distribution network is likely to improve the service levels, which will result in reducing delivery time and increasing customer satisfaction. To the extent that distribution costs constitute a large part of the total marketing costs for an organization, an updated network design can result in dollar savings and increased customer service. Second, and more important, distribution systems—if correctly designed—can not only offset marketing costs but also enable marketing expenditures to have a stronger effect on revenues generated from customers (e.g., by mitigating stockout costs). This is consistent with the achievement of a dual emphasis with marketing (Rust, Moorman, and Dickson 2002). The dual emphasis of this study is on achieving a lower total cost while increasing the customer service level and with an efficiency orientation advocated by leading researchers (Rust et al. 2004). Third, the results show the importance of using decision calculus to implement marketing strategy from the "inside out." In other words, managerial inputs about key decision variables can be used to design a system from the inside to obtain superior customer service and customer satisfaction. Such

FIGURE 6
Performance Comparison of Different Heuristics



an approach can be used by market-oriented firms even in the absence of direct customer inputs.

In terms of limitations, we acknowledge that different firms may have different constraints in their network design other than those addressed in our model. In addition, although inventory holding and backlogging costs were beyond the scope of this article under the assumptions of a deterministic model and a single period, they can be incorporated into the model for a more precise estimation of the distribution costs. Variants of this approach can be developed to suit the specific needs of a given firm. However, we hope that such developments will not only lead to additional theoretical insights about the importance of distribution network design but also spur research in marketing to enable a more thorough investigation of supply chain issues.

Appendix
List of Indexes, Parameters, and Variables in the CSCD Model

Indexes

i: Product type, where I is the total number of product types a company must transport (i = 1, 2, …, I).

j: Plant number, where J is the total number of plants existing in the supply chain (j = 1, 2, …, J).

k: RDC (warehouse) number, where K is the maximum number of RDCs that can be opened, which can be specified by management or set to a very large number by default. The optimal number of RDCs obtained will be equal to or less than the K value specified (k = 1, 2, …, K).

l: CDC number where L is the total number of CDCs to be allocated in the supply chain problem (l = 1, 2, …, L).

Parameters

$F(w_k)$: Cost function of opening and operating RDC k. It is a function of warehouse size.

g: Unit transportation cost from plant j to RDC k per weighted distance. The cost g can be derived from Equations 10, 12, and 14.

h: Unit transportation cost from RDC k to CDC l per weighted distance. The cost h can be derived from Equations 10, 11, and 13.

(a_j, b_j): Location coordinates of plant j.

(a_l', b_l'): Location coordinates of CDC l.

D_{il}: Demand for product i by CDC l.

Variables

(x_k, y_k): coordinates of RDC k.

$d_j(x_k, y_k)$: Distance from plant j to RDC k. $d_j(x_k, y_k) = \sqrt{(x_k - a_j)^2 + (y_k - b_j)^2}$.

$d_l(x_k, y_k)$: Distance from RDC k to CDC l. $d_l(x_k, y_k) = \sqrt{(x_k - a_l')^2 + (y_k - b_l')^2}$.

u_{kl}: Binary variable that takes the value of 1 if RDC k serves CDC l.

z_k: Binary variable that takes the value of 1 if RDC k is opened.

s_{ijk}: Amount of product i shipped from plant j to RDC k.

t_{ikl}: Amount of product i shipped from RDC k to CDC l.

w_k: Relative size of RDC k.

REFERENCES

Amiri, Ali (2006), "Designing a Distribution Network in a Supply Chain System: Formulation and Efficient Solution Procedure," *European Journal of Operational Research*, 171 (June), 567–76.

Anderberg, Michael R. (1973), *Cluster Analysis for Applications in Probability and Mathematical Statistics*. New York: Academic Press.

Anderson, Eric T., Gavan J. Fitzsimons, and Duncan Simester (2006), "Measuring and Mitigating the Costs of Stockouts," *Management Science*, 52 (November), 1751–63.

Arntzen, Bruce C., Gerald G. Brown, Terry P. Harrison, and Linda L. Trafton (1995), "Global Supply Chain Management at Digital Equipment Corporation," *Interfaces*, 25 (January–February), 69–93.

Ashlock, Daniel (2005), *Evolutionary Computation for Modeling and Optimization*. New York: Springer.

Baumol, William J. and Philip Wolfe (1958), "A Warehouse-Location Problem," *Operations Research*, 6 (March–April), 252–63.

Beamon, Benita M. (1998), "Supply Chain Design and Analysis: Models and Methods," *International Journal of Production Economics*, 55 (August), 281–94.

Brandeau, Margaret L. and Samuel S. Chiu (1989), "An Overview of Representative Problems in Location Research," *Management Science*, 35 (June), 645–74.

Brimberg, Jack, Pierre Hansen, Nenad Mladenovic, and Eric D. Taillard (2000), "Improvements and Comparison of Heuristics for Solving the Uncapacitated Multisource Weber Problem," *Operations Research*, 48 (May–June), 444–60.

Brown, Gerald G., Glenn W. Graves, and Maria D. Honczarenko (1987), "Design and Operation of a Multicommodity Production/Distribution System Using Primal Goal Decomposition," *Management Science*, 33 (November), 1469–80.

Bucklin, Louis P. (1967), "The Concept of Mass in Intra-Urban Shopping," *Journal of Marketing*, 31 (October), 37–42.

Chakravarti, Dipankar, Andrew Mitchell, and Richard Staelin (1979), "Judgment Based Marketing Decision Models: An Experimental Investigation of the Decision Calculus Approach," *Management Science*, 25 (March), 251–63.

Chan, Tat Y., V. Padmanabhan, and P.B. Seetharaman (2007), "An Econometric Model of Location and Pricing in the Gasoline Market," *Journal of Marketing Research*, 44 (November), 622–35.

Cohen, Morris A. and Hau L. Lee (1988), "Strategic Analysis of Integrated Production-Distribution Systems: Models and Methods," *Operations Research*, 36 (March–April), 216–28.

Cooper, Leon (1964), "Heuristic Methods for Location-Allocation Problems," *SIAM Review*, 6 (January), 37–53.

——— (1967), "Solutions of Generalized Locational Equilibrium Models," *Journal of Regional Science*, 7 (June), 1–19.

Cox, Reavis (1959), "Consumer Convenience and the Retail Structure of Cities," *Journal of Marketing*, 23 (April), 355–62.

Daskin, Mark S. (1995), *Network and Discrete Location: Models, Algorithms and Applications*. New York: John Wiley & Sons.

Davis, Peter (2006), "Spatial Competition in Retail Markets: Movie Theaters," *Rand Journal of Economics*, 37 (December), 964–82.

Deveci-Kocakoc, Ipek and Ali Sen (2006), "Utilising Surveys for Finding Improvement Areas for Customer Satisfaction Along the Supply Chain," *International Journal of Market Research*, 48 (5), 623–36.

Devlin, James F. and Philip Gerrard (2004), "Choice Criteria in Retail Banking: An Analysis of Trends," *Journal of Strategic Marketing*, 12 (March), 13–27.

Donthu, Naveen and Roland T. Rust (1989), "Estimating Geographic Customer Densities Using Kernel Density Estimation," *Marketing Science*, 8 (Spring), 191–203.

Du Merle, Olivier, Daniel Villeneuve, Jacques Desrosiers, and Pierre Hansen (1999), "Stabilized Column Generation," *Discrete Mathematics*, 194 (January), 229–37.

Eiben, Agostan E. and Jim E. Smith (2007), *Introduction to Evolutionary Computing*. New York: Springer.

Elhedhli, Samir and Jean-Louis Goffin (2005), "Efficient Production-Distribution System Design," *Management Science*, 51 (July), 1151–64.

Erenguc, Selcuk S., Natalie C. Simpson, and Asoo J. Vakharia (1999), "Integrated Production/Distribution Planning in Supply Chains: An Invited Review," *European Journal of Operational Research*, 115 (June), 219–36.

Erlenkotter, Donald (1978), "A Dual-Based Procedure for Uncapacitated Facility Location," *Operations Research*, 26 (November–December), 992–1009.

Eskigun, Erdem, Reha Uzsoy, Paul V. Preckel, George Beaujon, Subramanian Krishnan, and Jeffrey D. Tew (2005), "Outbound Supply Chain Network Design with Mode Selection, Lead Times and Capacitated Vehicle Distribution Centers," *European Journal of Operational Research*, 165 (August), 182–206.

Frontline Systems (2007), *Premium Solver*. Incline Village, NV: Frontline Systems.

Geoffrion, Arthur M. and Glenn W. Graves (1974), "Multicommodity Distribution System Design by Benders Decomposition," *Management Science*, 20 (January), 822–44.

——— and Richard F. Powers (1995), "20 Years of Strategic Distribution System Design: An Evolutionary Perspective," *Interfaces*, 25 (September–October), 105–127.

Ghosh, Avijit and C. Samuel Craig (1983), "Formulating Retail Location Strategy in a Changing Environment," *Journal of Marketing*, 47 (Summer), 56–68.

——— and ——— (1986), "An Approach to Determining Optimal Locations for New Services," *Journal of Marketing Research*, 23 (November), 354–62.

——— and Sara L. McLafferty (1982), "Locating Stores in Uncertain Environments: A Scenario Planning Approach," *Journal of Retailing*, 58 (Winter), 5–22.

Gower, John C. (1985), "Measures of Similarity, Dissimilarity and Distance" in *Encyclopedia of Statistical Sciences*, S. Kotz, Normal L. Johnson, and Campbell B. Read, eds. New York: John Wiley & Sons, 397–405.

Holland, John H. (1975), *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.

Iyer, Ananth V. and Mark E. Bergen (1997), "Quick Response in Manufacturer-Retailer Channels," *Management Science*, 43 (April), 559–70.

Jayaraman, Vaidyanathan and Arthur Ross (2003), "A Simulated Annealing Methodology to Distribution Network Design and Management," *European Journal of Operational Research*, 144 (February), 629–45.

Kadiyali, Vrinda, Pradeep Chintagunta, and Naufel Vilcassim (2000), "Manufacturer-Retailer Channel Interactions and Implications for Channel Power: An Empirical Investigation of Pricing in a Local Market," *Marketing Science*, 19 (Spring), 127–48.

Klamroth, Kathrin (2001), "Planar Weber Location Problems with Line Barriers," *Optimization*, 49 (5–6), 517–27.

Klose, Andreas and Andreas Drexl (2005), "Facility Location Models for Distribution System Design," *European Journal of Operational Research*, 162 (April), 4–29.

Korpela, Jukka and Antti Lehmusvaara (1999), "A Customer Oriented Approach to Warehouse Network Evaluation and Design," *International Journal of Production Economics*, 59 (March), 135–46.

Kuo, Ren J., Sheng-Chai Chi, and Shan-Shin Kao (2002), "A Decision Support System for Selecting Convenience Store Location Through Integration of Fuzzy AHP and Artificial Neural Network," *Computers in Industry*, 47 (February), 199–214.

Langston, Paul, Graham P. Clarke, and Dave B. Clarke (1997), "Retail Saturation, Retail Location, and Retail Competition: An Analysis of British Grocery Retailing," *Environment and Planning: A*, 29 (January), 77–104.

Little, John D.C. (2004), "Comments on 'Models and Managers: The Concept of a Decision Calculus,'" *Management Science*, 50 (December), 1854–60.

Lord, Dennis J. and Charles Lynds (1981), "The Use of Regression Models in Store Location Research: A Review and Case Study," *Akron Business and Economic Review*, 12 (Summer), 13–19.

Mahajan,Vijay, Subhash Sharma, and Roger A. Kerin (1988), "Assessing Market Opportunities and Saturation Potential for Multi-Store, Multi-Market Retailers," *Journal of Retailing*, 64 (Fall), 315–32.

Meepetchdee, Yongyut and Nilay Shah (2007), "Logistical Network Design with Robustness and Complexity Considerations," *International Journal of Physical Distribution & Logistical Management*, 37 (March), 202–222.

Melkote, Sanjay and Mark S. Daskin (2001), "Capacitated Facility Location/Network Design Problems," *European Journal of Operational Research*, 129 (March), 481–95.

Menon, Anil (2004), *Frontiers of Evolutionary Computation* (*Genetic Algorithms and Evolutionary Computation*). New York: Springer Publishing Company.

Mittal, Vikas, Wagner A. Kamakura, and Rahul Govind (2004), "Geographic Patterns in Customer Service and Satisfaction: An Empirical Investigation," *Journal of Marketing*, 68 (July), 48–62.

Moon, Douglas I. and Sohail S. Chaudhry (1984), "An Analysis of Network Location Problems with Distance Constraints," *Management Science*, 30 (March), 290–307.

Mulhern, Francis J. (1997), "Retail Marketing: From Distribution to Integration," *International Journal of Research in Marketing*, 14 (May), 103–124.

Murry, John P., Jr., and Jan B. Heide (1998), "Managing Promotion Program Participation Within Manufacturer–Retailer Relationships," *Journal of Marketing*, 62 (January), 58–68.

Pinkse, Joris, Margaret E. Slade, and Craig Brett (2002), "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70 (May), 1111–53.

Rosing, Kenneth E. (1992), "An Optimal Method for Solving the (Generalized) Multi-Weber Problem," *European Journal of Operational Research*, 58 (May), 414–26.

Rust, Roland T., Tim Ambler, Gregory S. Carpenter, V. Kumar, and Rajendra K. Srivastava (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (October), 76–89.

——— and Naveen Donthu (1995), "Capturing Geographically Localized Misspecification Error in Retail Store Choice Models," *Journal of Marketing Research*, 32 (February), 103–110.

———, Christine Moorman, and Peter R. Dickson (2002), "Getting Return on Quality: Revenue Expansion, Cost Reduction, or Both?" *Journal of Marketing*, 66 (October), 7–24.

Saourirajan, Karthik, Leyla Ozsen, and Reha Uzsoy (2007), "A Single-Product Network Design Model with Lead Time and

Safety Stock Considerations," *IIE Transactions*, 39 (May), 411–24.

Shen, Zuo-Jun M. (2005), "A Multi-Commodity Supply Chain Design Problem," *IIE Transactions*, 37 (August), 753–62.

Swersey, Arthur and Lakshman S. Thakur (1995), "An Integer Programming Model for Locating Vehicle Emission Testing Stations," *Management Science*, 41 (March), 496–512.

Thelen, Eva M. and Arch G. Woodside (1997), "What Evokes the Brand or Store? Consumer Research on Accessibility Theory Applied to Modeling Primary Choice," *International Journal of Research in Marketing*, 14 (May), 125–45.

Vidal, Carlos J. and Marc Goetschalckx (1997), "Strategic Production-Distribution Models: A Critical Review with Emphasizes on Global Supply Chain Models," *European Journal of Operational Research*, 98 (April), 1–18.

Wen, Meilin and Kakuzo Iwamura (2008), "Fuzzy Facility Location-Allocation Problem Under the Hurwicz Criterion," *European Journal of Operational Research*, 184 (January), 627–35.

Wesolowsky, George O. (1993), "The Weber Problem: History and Perspectives," *Location Science*, 1 (January), 5–23.