Consolidating a Warehouse Network: A Physical Programming Approach

Emanuel Melachrinoudis

Achille Messac

Hokey Min

Corresponding Author

Achille Messac, Ph.D.
Distinguished Professor and Department Chair
Mechanical and Aerospace Engineering
Syracuse University, 263 Link Hall
Syracuse, New York 13244, USA

Email: messac@syr.edu
Tel: (315) 443-2341
Fax: (315) 443-3099
https://messac.expressions.syr.edu/

Bibliographical Information

Consolidating a warehouse network: A physical programming approach

Emanuel Melachrinoudis\textsuperscript{a,*}, Achille Messac\textsuperscript{b}, Hokey Min\textsuperscript{c}

\textsuperscript{a}Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115, USA
\textsuperscript{b}Department of Mechanical, Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
\textsuperscript{c}UPS Center for World Wide Supply Chain Management, University of Louisville, Louisville, KY 40292, USA

Received 1 June 2003; accepted 1 April 2004
Available online 22 July 2004

Abstract

Faced with mounting cost pressure, many companies consider downsizing their distribution networks in ways that involve consolidation or phase-out of some of their existing warehousing facilities. To effectively reconfigure a warehouse network through consolidation and elimination, this paper proposes a novel multiple criteria methodology called physical programming (PP). The proposed PP model enables a decision maker to consider multiple criteria (i.e., cost, customer service and intangible benefits) and to express criteria preferences not in a traditional form of weights, but in ranges of different degrees of desirability. The proposed model is tested with real data involving the reconfiguration of an actual company’s distribution network in the United States and Canada.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Network design; Supply chain; Physical programming; Multiple criteria decision making

1. Introduction

Faced with increasing competition and mounting cost pressures, a growing number of firms are considering re-configuring or re-engineering their corporate structures through the consolidation and phase-out of some of their existing warehouses. In fact, the recent surveys conducted by the Warehouse Education Research Council (e.g. WERC, 1999, 2002) predicted a gradual decline in the total number of warehouses that caused an increase in their average size throughout the supply chain.

Such a trend can be illustrated by the following two cases, one from the industry and another from the Government:

(1) Gillette embarked on a very ambitious re-structuring program by closing several factories and 13 distribution centers worldwide,
which is intended to generate savings of more than $125 million annually on an ongoing basis (e.g. Johnson, 2001). Specific program activities included streamlining the supply chain via warehouse consolidation, followed by the downsizing and centralization of corporate functions.

(2) In addition to providing all fuel, medical care, subsistence and clothing support, the defense logistics support command (DLSC) of the Department of Defense is responsible for procuring, storing, and distributing 85% of the consumable spare parts required by US military services. The worldwide logistics network encompasses the Defense Distribution Center, 5 inventory control points, and 24 distribution depots. DLSC has dramatically reduced its infrastructure by streamlining business practices and downsizing excess capacity. The 27% reduction in warehouse space resulted in a 29% drop in personnel over a 5-year period, representing millions of dollars of savings in military spending and a subsequent decrease in the financial burden of the American taxpayers (e.g. Morton, 1999).

As illustrated above, the consolidation of warehouses can help a company save transportation, inventory and warehousing costs due to economies of scale. To elaborate, the square root rule shows that the reduced number of warehouses can decrease total inventory carrying costs (e.g. Ballou, 1999). Also, the lower variance of the aggregate demand can reduce the chance of stockouts (e.g. Bordley et al., 1999). Total transportation costs can be reduced, due to increased opportunities for large-volume shipments and the subsequent negotiation leverage for better freight rates. Total transportation costs can be further reduced by eliminating cross-hauling among too many warehouses. Furthermore, material handling costs would decrease due to increased opportunities for bulk storage and mass picking at centralized locations. Central administrative costs can be lessened through managing fewer warehouses.

Despite its various cost saving potentials, warehouse consolidation has a drawback. It lengthens lead-time, and consequently may deteriorate customer service. This is because the fewer the warehouses, the longer the distances from customers. To maintain an acceptable level of customer service, many companies want to offer next-day delivery services to the majority of their customers. Considering hours of service regulations, stipulated by the US Federal Highway Administration (FHA), a majority of the company’s customers should be within 10 hours of driving time from its warehouses.

Other factors affecting warehouse network reconfiguration are labor-management relations, labor quality (e.g. skill and educational level of the available workforce in the potential site), tax incentives and market potentials. The high risk of labor strikes or employee turnover is a key concern in warehousing operations because such a risk can disrupt the smooth flow of products to customers (e.g. Ackerman, 1994). Thus, a company may wish to consolidate a warehouse at a site where good labor-management relations exist and the quality of life can help the company retain a highly skilled workforce. On the contrary, a company may wish to close or relocate a warehouse with a history of strikes or high employee turnover. In their efforts to attract business, many states today provide tax refunds, job training grants and direct grants for facility relocation/expansion (e.g. Bowman, 1998). Indeed, labor quality, tax incentives, and loans are considered the most important location factors by thousands of economic development executives (e.g. Venable, 1996). Another factor in favor of maintaining a warehouse, or even expanding it, is the potential for increasing a market share. In a broad sense, the market potential of a trading area is dictated by demographics, the business climate, and level of competition. These intangible benefits of warehouse sites should therefore be part of the criteria in order to best determine which warehouses to keep open and which to phase-out or consolidate with others.

The literature on the location of production plants and warehouses is extensive and broad in its scope (see, for example, Meidan (1978), Krarup and Pruzan (1983), Aikens (1985), and Geoffrion et al. (1995) for excellent reviews of the related
literature). Most of the existing models, however, are not designed to deal with dynamic changes in parameters. For instance, a few years after a distribution network has been in operation, customer bases may have shifted, warehousing and transportation costs may have changed, and customer expectations for better service may have risen. Thus, the need for periodic re-evaluation and reconfiguration of a distribution network is paramount for the survival and viability of a company.

Despite such a need, the literature on restructuring a distribution network is very sparse. Melachrinoudis et al. (2000) developed a multi-criteria model for relocating a single plant/warehouse facility by assessing the impact of relocation on the whole supply chain, including suppliers, manufacturing plants, warehouses and customers. Watts (2000) used a personal computer-based center of gravity model to solve the problem of consolidating regional chemical distribution facilities in the Midwest and Eastern United States. Min and Melachrinoudis (2001) outlined various strategies and models for restructuring a warehouse network. They modified a traditional \( p \)-median model to dramatically reduce the number of warehouses of a major manufacturer and distributor of plastic films in the United States and Canada. In this paper, the same authors present a more comprehensive multi-criteria model for consolidating a warehouse network and validating it using the company’s data.

Several authors have recognized the benefits of warehouse centralization versus decentralization. Patton (1986) summarized the benefits of warehouse centralization: reduced inbound shipping costs, improved inventory management, reduced safety stock and better opportunity for negotiating transportation services. On the other hand, he listed some benefits of warehouse decentralization: rapid response to customers and reduction in outbound shipping costs. Similarly, Chopra (2003) observed that as the number of warehouses increased, inventory and facility costs increased, whereas, inbound transportation costs increased and outbound transportation costs decreased. Some authors addressed the question of how to achieve centralization through the consolidation of inventories. For example, Ballou (1981) used a regression analysis model for estimating aggregate safety stock of consolidated stocking points by using actual data. His model revealed that the safety stock aggregation differed from industry to industry. Das (1978) proposed a dynamic programming approach to allocate inventory over a number of locations and compared the total cost of centralized versus decentralized inventories.

Although the aforementioned models identified certain benefits of warehouse consolidation, there is no mathematical model in the literature that took into account all factors affecting the consolidation of a distribution network. To fill such a void, we developed a multiple objective model for the Consolidation of a Warehouse Network. Three criteria are considered: (1) minimization of total distribution costs comprised of production costs, fixed and variable warehousing costs, warehouse relocation costs, inbound and outbound distribution costs; (2) maximization of customer services that can be rendered to customers in terms of acceptable delivery time (coverage), and (3) maximization of intangible benefits associated with the new distribution network.

The model is structured within a novel multi-objective optimization framework, linear physical programming (LPP). To clarify the link between physical programming (PP) and existing methods, a brief review of multiobjective methods is undertaken in Section 2. LPP is described in Section 3. The model is presented in Section 4 and it is tested and validated in Section 5 using actual data obtained from a company that considered consolidating its distribution network. The paper ends with a summary and conclusions in Section 6.

2. Multiobjective methods

There are many methods in the literature for solving multiobjective models. We present a brief synopsis of multiobjective methods for the purpose of establishing the relationship between the proposed method, PP, and existing ones. As Ignizio and Cavalier (1994) note, solution methods are based upon certain philosophies and there is no “one right way” to approach a problem...
involving multiple, conflicting objectives. Some methods require the decision maker (DM) to articulate his/her preferences early on in the decision-making process, others require a progressive articulation of DM’s preferences, and yet others do not require any input from the DM. Therefore, an important characteristic of solution methods is the amount and type of preference information requested from the DM. Another important classification of multiobjective methods is according to the solution space. Certain methods are more suitable for problems with a discrete number of alternatives while others apply to problems exhibiting continuous solution spaces.

Among the discrete methods, also called multi-criteria analysis (MCA) methods, the best known are (a) the multiattribute utility theory (e.g. Keeney and Raiffa, 1976), (b) the analytic hierarchy process (AHP) (e.g. Saaty, 1988) and (c) the outranking methods (e.g. Benayoun et al., 1966; Roy, 1971). All these methods involve implicit or explicit aggregation of each alternative’s preference across all objectives (criteria) to form an overall assessment of each alternative, on the basis of which the set of alternatives can be compared. The principal difference between the methods is the way in which this aggregation is performed.

Multiattribute utility is an elegant axiomatic theory that can be used in making decisions under uncertainty. The preference information is elicited from the DM in an elaborate process using lotteries to build individual utility functions of the objectives (attributes). Although well regarded and effective, in its most general form, the method is relatively complex in that it demands significant time and expertise from the DM. When certain conditions are satisfied however the aggregate objective function can be reduced to two simple forms, the additive (linear) function and the multiplicative function of the individual utility functions of the objectives.

An AHP develops a linear aggregate function in a less sophisticated way. The distinctive feature of the method is the structuring of the objectives and alternatives (elements) in a tree hierarchy. The DM’s preferences are elicited in the form of pairwise comparisons among the elements of each level in the hierarchy using a 1–9 scale that expresses the importance or dominance of one element over another. A strength of this method is that generally DMs find the pairwise comparison form of preference information straightforward and convenient. On the other hand, doubts have been raised about the method’s theoretical foundation and about the possibility of occurrence of the rank reversal phenomenon.

A rather different approach from the previous two methods is followed by the outranking methods. They use the so-called outranking relationship to eliminate alternatives and thus reduce the original set to a subset of alternatives. In simple terms, an alternative outranks another if it outperforms the other on enough objectives of sufficient importance (as reflected by the sum of criteria weights) and is not outperformed by the other alternative in the sense of exhibiting inferior performance on any one objective. All alternatives are then evaluated in terms of the extent to which they exhibit sufficient outranking as measured against a pair of threshold parameters. The main concern about this method is that it is dependent on some rather arbitrary definitions of what constitutes outranking and how threshold parameters are set and later manipulated by the DM.

Among the continuous methods, also called multi-criteria optimization (MCO) methods, the best known are the (a) generating method (e.g. Steuer, 1986), (b) interactive methods (e.g. Cohon, 1978; Hwang and Masud, 1979), and (c) goal programming methods (e.g. Lee, 1972; Ignizio, 1976) including fuzzy programming (e.g. Zimmermann, 1978, 1985). The generating method or vector-maximum approach generates all efficient extreme points of a model from which the DM selects his/her most preferred one. It does not require any input from the DM. Although the method’s theoretical foundation is solid and the idea of efficiency is undisputable, the method has some drawbacks. Among the problems associated with this method are its high computational requirements to compute all efficient solutions and the difficulty of the DM to sort out the large number of generated solutions for the most-preferred one.
The interactive methods require a progressive articulation of preferences from the DM, i.e., they involve extensively the DM throughout the solution process. An efficient solution is first identified, followed by the DM’s expression of his/her preference usually in terms of tradeoffs, which in turn defines a search direction for another solution. These steps are repeated until an acceptable solution is reached. Although most researchers agree that these methods have a clear advantage among MCO methods, because they actively involve the DM in the decision-making process, these methods have not been used significantly in practice. The reason for these methods limited use is that, to be effective, the DM needs to be very unusual—one who is willing to commit a great deal of time and effort to the (technical aspects of the) solution process (e.g. Ignizio and Cavalier, 1994).

The distinctive feature of goal programming (GP) methods is the establishment of a goal level of achievement for each objective. A solution is sought that achieves the goals “as closely as possible”. Depending how closeness is defined, three most popular (as well as practical) forms of GP have been developed: (a) the Archimedian GP, (b) the lexicographic GP, and (c) the Chebyshev GP. All of them minimize some function of deviations from the goals, the penalty function. The Archimedian GP minimizes the sum of the weighted deviations from the goals; the lexicographic GP prioritizes the goals and minimizes deviations according to priorities starting with the highest priority goal; the Chebyshev GP minimizes the maximum (minimax) deviation over all goals.

Another multiobjective method that was developed independently of GP, fuzzy programming, appears to be very similar to Chebyshev GP. Each objective function value of any given feasible solution can be considered a fuzzy number between the best and worst values of that objective. The latter are obtained by maximizing and minimizing each objective independently subject to the original set of constraints. The best objective values are considered as goal levels and the range of each objective is computed as the difference between the maximum and minimum value. For a given feasible solution, the fuzzy membership function value of an objective is defined as the deviation from the goal divided by that objective’s range, thus a value between 0 and 1. Fuzzy programming finds the solution that minimizes the largest fuzzy membership function. Thus, it is identical to Chebyshev GP, except that the maximum deviation is weighted by the ranges in the goal constraints (e.g. Ignizio and Cavalier, 1994). In addition to the goals, other input parameters can be represented by fuzzy numbers and fuzzy arithmetic can be used to combine fuzzy numbers based on fuzzy set theory (e.g. Zadeh, 1965). Thus, fuzzy extensions to conventional MCA methods have been developed, such as fuzzy utility theory, fuzzy AHP and fuzzy outranking methods. Fuzzy programming has enjoyed popularity because it responds to the imprecision that surrounds much of the data, including goals, in multiobjective problems and the explicit way of representing that vagueness via fuzzy sets.

Goal programming penalizes small and large deviations from a given goal alike through the same weight. To offset that criticism, Charnes and Cooper (1977) devised a more elaborate means of penalizing deviations by defining multiple achievement levels. The resulting penalty function is piecewise linear with multiple weights. Although more accurate, the new version of Archimedian goal programming has not been extensively applied due to the inherent difficulty in determining additional weights. In fact, a general criticism of several multiobjective methods is that the DM has a great difficulty in determining weights, therefore in many situations is unwilling to specify weights (e.g. Lootsma, 1997) or the weights specified may not reflect accurately his/her preference function. A method that circumvents the weight difficulty while maintaining the benefits of Charnes and Cooper representation of each objective through the multiple levels of desirability is PP. PP captures the idea that our natural language in discussing achievement or performance is not precise, much like what the fuzzy sets approach advocates. For example, the DM specifies the ideal range of an objective, the desirable range of an objective, the tolerable range, and so on. Next, we provide a synopsis of the PP method.
3. Linear physical programming

LPP is a multiobjective optimization method that develops an aggregate objective function of the criteria in a piecewise Archimedian goal programming fashion. The PP approach in its nonlinear (general) form was developed by Messac (1996) and it has been used primarily in engineering design optimization; see, for example, Messac and Hattis (1996), Messac and Wilson (1998), Messac (1998, 2000), Messac et al. (1999) and Messac and Ismail-Yahaya (2002). In its piecewise linear form, LPP has been used in several diverse applications. Maria et al. (2003) used LPP in production planning. Melachrinoudis et al. (2000) used LPP to optimize relocation decisions for a manufacturing/distribution facility. This section provides a brief outline of LPP. For a comprehensive presentation, see Messac et al. (1996).

Within the LPP procedure, the DM expresses his/her preference with respect to each criterion using four classes, i.e., the DM declares each criterion as belonging to one of four distinct classes. Fig. 1 depicts the qualitative and quantitative meaning of each class. Let \( x \) be the decision vector and \( g_p(x) \) be the \( p \)th generic linear objective function. The value of this objective function is on the horizontal axis, and the penalty function that is minimized for that criterion, \( z_p \), is on the vertical axis. A lower value of the penalty function is better than (more valuable than) a higher value thereof. The ideal value of the penalty function is zero. The penalty functions become additive constituent components of the aggregate objective function (to be minimized). A criterion falls into one of four classes of penalty functions, hereby called class functions, defined as follows:

- **Class 1S**: smaller-is-better, i.e., minimization
- **Class 2S**: larger-is-better, i.e., maximization
- **Class 3S**: value-is-better
- **Class 4S**: range-is-better

The structure of the above class functions provides the means for DMs to express the ranges of differing levels of preference for each criterion.

The PP lexicon comprises terms that characterize the degrees of desirability of each criterion, using six types of ranges. For Class 1S, shown in Fig. 1, the ranges are defined as follows:

- **Ideal** range \((g_p \leq t_{p1}^{+} \text{; Range-1})\): A range over which every value of the criterion is ideal (the most desirable possible). Any two points in that range are equally desirable to the DM.
- **Desirable** range \((t_{p1}^{+} \leq g_p \leq t_{p2}^{+} \text{; Range-2})\): An acceptable range that is desirable.
- **Tolerable** range \((t_{p2}^{+} \leq g_p \leq t_{p3}^{+} \text{; Range-3})\): An acceptable, tolerable range.
- **Undesirable** range \((t_{p3}^{+} \leq g_p \leq t_{p4}^{+} \text{; Range-4})\): A range that, while acceptable, is undesirable.

Fig. 1. LPP class functions.
Highly undesirable range ($r_{ps}^- < g_p < r_{ps}^+$; Range-5): A range that, while still acceptable, is highly undesirable.

Unacceptable range ($g_p > r_{ps}^+$; Range-6): The range of values that the criterion may not take.

The range-boundaries (target levels), $r_{ps}^+$ through $r_{ps}^-$, are physically meaningful values that are chosen by the DM (company executive) to quantify the preference associated with the $p$th criterion. The ranges for the other classes and the associated target levels are defined in a like manner. Consider for example the customer coverage objective in this study. The DM expresses his preference regarding the percentage of customer demand that can be delivered from warehouses to customers within 10 hours or less as follows: (1) ideal, at least 97%; (2) desirable, 90–97%; (3) tolerable, 80–90%; (4) undesirable, 70–80%; (5) highly undesirable, 50–70%; and (6) unacceptable, lower than 50%. These ranges are set based on the industry norms and the company’s past service performances. LPP uses the preceding information—and similar statements regarding other criteria—to determine incremental criteria weights and form the aggregate objective function. Thus, LPP removes the necessity to choose weights, which is undoubtedly a difficult task. LPP allows the DM to reflect his preferences for each criterion in an explicit and flexible manner.

Certain desirable properties of the class functions are used to provide the theoretical framework for their computation. These properties are summarized in the appendix. Based on the properties of class functions, LPP determines the weights ($\tilde{w}_{ps}^+$ and $\tilde{w}_{ps}^-$) that represent the incremental slopes of the penalty functions, $z_p$, thus enabling $z_p$ to be expressed as a piecewise linear function of criterion $g_p$. The aggregate objective function (to be minimized) is then constructed as a weighted sum of deviations over all ranges ($s = 2, \ldots, 5$) and criteria ($p = 1, \ldots, P$). The resulting LPP formulation is as follows:

$$\min_{d_{ps},d_{ps}^-} J = \sum_{p=1}^{P} \sum_{s=2}^{5} \left( \tilde{w}_{ps}^- d_{ps}^- + \tilde{w}_{ps}^+ d_{ps}^+ \right)$$

subject to:

$$g_p(x) - d_{ps}^- \leq r_{ps}^{+1} - g_p \leq d_{ps}^+; \quad g_p(x) \leq r_{ps}^+$$

(for all classes 1S, 3S, 4S; $p = 1, \ldots, P$; $s = 2, \ldots, 5)$$

$$g_p(x) + d_{ps}^- \geq r_{ps}^{-1}; \quad d_{ps}^- \leq 0; \quad g_p(x) \geq r_{ps}^-$$

(for all classes 2S, 3S, 4S; $p = 1, \ldots, P$; $s = 2, \ldots, 5)$$

$x \in S$.

The variables $d_{ps}$ and $d_{ps}^-$ in the above formulation are the negative and positive deviations of the objective function (goal) value $g_p(x)$ from target levels $r_{ps}^{+(1)}$ and $r_{ps}^{-1}$, respectively, and $g_p(x)$ is a linear function of $x$. Following the Archimedian objective function (1) are two types of goal constraints. Goal constraints (2) apply to criteria belonging to all classes but Class 2S, while goal constraints (3) apply to criteria belonging to all classes except Class 1S. Finally, constraints (4) define the feasible region $S$ of the decision vector, $x$; i.e. they are the system constraints.

It is important at this point to make some important observations. Regarding the specification of desirability ranges, rather than that of the typical weights for each objective, the reader is encouraged to read the papers that develop the method (e.g. Messac, 1996; Messac et al., 1996). We summarize here by saying that the process of changing the preference ranges is physically meaningful, unlike changing the weights. If the desirable cost is changed from 15 to 10, it is very clear what one is doing by changing the associated preference numbers. However, the corresponding weights might involve a change from 753 to 1053, which are both void of physical meaning. The only thing that one knows is that the weight needed increase, but not its magnitude. Further, assigning one weight for each criterion fails to reflect the different desirability level for different regions. To do so, the ability to form a preference function of the type of Fig. 1 is required. Unfortunately, the notorious difficulty that would entail determining four to eight weights (slopes) for each criterion is well known. LPP indeed provides (i) the means to
offer the flexibility of Fig. 1, (ii) the ability to directly define preference in terms of the values of the criteria explicitly, and (iii) an algorithm that automatically defines the weights (slopes) for use in the model (the DM does not need to guess what these slopes should be).

In addition, we note that the linear LPP method provides for a discrete set of weights (slopes). This approach is in general adequate for problems that are formulated in the linear programming domain. Using the linear LPP model allows the process to fully take place in the linear programming domain. However, if for any reason it is deemed desirable to have a continuous change rather than a discrete set of changes, then the nonlinear PP model can be used. However, the associated complication is that the immense benefit of linear programming will no longer apply. The authors advise that when the problem model is linear, the linear LPP model be used.

4. Model design

Prior to developing the mathematical model, we make the following underlying assumptions:

- A company produces and distributes a single product.
- Existing warehouse capacities are given in terms of maximum annual throughput (volume of products). No partial capacity is consolidated into another warehouse. That is, either a warehouse is closed (and its capacity lost), or its whole capacity is relocated and consolidated into another existing warehouse.

Considering the multitude of conflicting objectives affecting the reconfiguration of a distribution network and the types of decisions to make, as outlined in Section 1, we formulate the warehouse consolidation problem as a multi-criteria mixed-integer programming model. The first objective is to minimize annual costs, or equivalently, to maximize annual potential cost savings due to phase-out and consolidation. The second objective is to maximize customer service by maximizing customer coverage within a stipulated distance (i.e., ten hours of driving distance). The other factors, outlined in the previous section, constitute the intangible benefits objective to be maximized. The proposed model is designed to determine which warehouses to close, which to retain, and which to consolidate with others. In addition, the model properly reassigns customers to warehouses in the new distribution network, and reallocates the plants’ production to the various warehouses.

The mathematical notation and formulation are as follows:

Indices: 
- \( n \) is an index for manufacturing plants \((n \in N)\).
- \( k \) is an index for customers \((k \in K)\).
- \( i \) is an index for existing warehouses \((i \in I)\).

Model variables:
- \( y_{ni} \) is the volume of products supplied by plant \( n \) to warehouse \( i \).
- \( x_{ik} \) is the volume of products shipped from warehouse \( i \) to customer \( k \).

\[
\begin{align*}
\forall i \neq j, \quad & z_{ji} = \begin{cases} 
1 & \text{if capacity of warehouse } j \text{ is relocated to site } i, \\
0 & \text{otherwise.}
\end{cases} \\
& \text{if existing warehouse } j \text{ remains open}
\end{align*}
\]

The negative and positive deviations from range targets are \( d_{ps}^- \) and \( d_{ps}^+ \) \((p = 1, \ldots, 3, s = 2, \ldots, 5)\), respectively.

Objectives:
- \( g_1 \) is the total annual cost (in millions of dollars).
- \( g_2 \) is the total customer demand (in %) that can be delivered within the stipulated access time \( \tau \).
- \( g_3 \) is the aggregate intangible benefits of all warehouses weighted by the proportion of demand they serve (aggregate score between 0 and 100).

Model parameters:
- \( c_i \) is the throughput capacity of existing warehouse \( i \).
- \( b_k \) is the demand of customer \( k \).
- \( f_i^c \) is the cost per unit capacity of warehouse \( i \).
- \( f_i^{cm} \) is the fixed cost of maintaining warehouse \( i \) (excluding capacity cost).
- \( f_i^s \) is the cost savings resulting from the closure of existing warehouse \( i \).
- \( l_i \) is the extent of intangible local incentives for warehouse \( i \) (in scores ranging from 0 to 100).
- \( q_n \) is the production capacity of manufacturing plant \( n \).
- \( r_{ji} \) is the cost of moving and relocating unit capacity from warehouse \( j \) to centralized (consolidated) site \( i \) \((j \neq i)\).
- \( s_{ik} \) is the cost of warehousing a unit product at warehouse \( i \) (based on average warehousing time) and shipping...
it from warehouse $i$ to customer $k$. $t^+_i$ and $t^-_i$ are the range targets ($p = 1, \ldots, 3$, $s = 1, \ldots, 5$). $h_{ik}$ is the delivery time (in hours) from warehouse $i$ to customer $k$. $\tau$ is the maximum allowable delivery time (hours) from warehouses to customers. $v_{ni}$ is the cost of producing and warehousing a unit of product at manufacturing plant $n$ plus shipping a unit product from manufacturing plant $n$ to warehouse $i$. $C(i)$ is the set of customers that can be reached from warehouse $i$ in $\tau$ hours, or $C(i) = \{ k | h_{ik} \leq \tau \}$. $\tilde{w}^+_p$ and $\tilde{w}^-_p$ are the incremental weights associated with negative and positive deviations from range targets ($p = 1, \ldots, 3$, $s = 2, \ldots, 5$). These weights are determined by LPP. The formulation equations are presented next, and a description of these equations immediately follows.

### 4.1. Physical programming formulation

1. Piecewise Archimedian aggregate function:

Minimize $J = \sum_{s=2}^{5} (\tilde{w}^+_s d^+_i + \tilde{w}^-_s d^-_i + \tilde{w}^+_a d^-_i)$ \hspace{1cm} (5)

subject to:

2. Criteria

$$
\begin{align*}
    g_1 &= \sum_{n \in N} \sum_{i \in I} v_{ni} y_{ni} + \sum_{i \in I} \sum_{k \in K} s_{ik} x_{ik} + \sum_{j \in I} \sum_{i \in I} t_{ji} z_{ji} \\
    &\quad + \sum_{j \in I} \sum_{i \in I} c_{ji} z_{ji} + \sum_{j \in I} f_j^m z_{ji} \\
    &\quad - \sum_{j \in I} \left( f_j^1 \left( 1 - \sum_{i \in I} z_{ji} \right) + f_j^m \sum_{i \in I, i \neq j} z_{ji} \right), \\
    g_2 &= 100 \left( \sum_{i \in I} \sum_{k \in K} x_{ik} \right) / \left( \sum_{k \in K} b_k \right), \\
    g_3 &= \left( \sum_{i \in I} \sum_{k \in K} x_{ik} \right) / \left( \sum_{k \in K} b_k \right).
\end{align*}
$$

3. Goal constraints

$$
\begin{align*}
    g_1 &\leq t_{i,s}, \\
    g_2 &\leq t_{i,s}, \\
    g_3 &\geq t_{i,s}, \\
    d^+_i, d^-_i, d^+_a &\geq 0, \quad s = 2, \ldots, 5.
\end{align*}
$$

4. System constraints

$$
\begin{align*}
    \sum_{i \in I} y_{ni} &\leq q_n \quad \forall n \in N, \\
    \sum_{n \in N} y_{ni} &= \sum_{k \in K} x_{ik} \quad \forall i \in I, \\
    \sum_{k \in K} x_{ik} &\leq \sum_{j \in I} c_{ji} z_{ji} \quad \forall i \in I, \\
    \sum_{i \in I} x_{ik} &= b_k \quad \forall k \in K, \\
    \sum_{z_{ji} \in [I]} z_{ji} &\leq 1 \quad \forall j \in I, \\
    y_{ni} &\geq 0 \quad \forall n \in N, \ i \in I, \\
    x_{ik} &\geq 0 \quad \forall i \in I, \ k \in K, \\
    z_{ji} &\in (0, 1) \quad \forall j, \ i \in I.
\end{align*}
$$

The formulation consists of four parts: the aggregate objective function, the criteria, the goal constraints and the system constraints. The aggregate objective function (5) minimizes the weighted sum of deviations from all range targets. Expressions (6)–(8) define the three criteria: the first (cost) is to be minimized, while the second (demand coverage) and the third (intangible benefits) are to be maximized. Each criterion gives
rise to five goal constraints, one for each range target. Criterion (6) generates constraints (9) and (10), criterion (7) generates constraints (11) and (12), and criterion (8) generates constraints (13) and (14). Constraint (15) restricts all deviational variables to nonnegative values.

Part 4 of the mathematical formulation contains the system (ordinary) constraints. Constraint (16) ensures that the total volume of products shipped to warehouses does not exceed the capacity of a manufacturing plant to supply such products. Constraint (17) ensures that the total volume of products shipped from all the manufacturing plants to each warehouse equals to the total volume of products shipped from that warehouse to its customers. Constraint (18) requires that the total volume of products shipped to customers does not exceed the throughput capacity of the serving warehouse. Constraint (19) ensures that each customer’s demand is satisfied. Constraint (20) states that the current capacity of an existing warehouse cannot be consolidated into another existing warehouse, unless such a consolidated warehouse remains open. Herein, |I| stands for cardinality of set I. Constraint (21) specifies status options for an existing warehouse j in the new network. It remains open \((z_{jj} = 1)\), or its capacity is consolidated into another existing warehouse \(i \in I, \ i \neq j\), \((z_{ji} = 1)\), or it is closed \((z_{ji} = 0)\). Constraints (22) and (23) restrict decision variables, \(y_{ji}\), and \(x_{ik}\) to nonnegative values, respectively, while constraint (24) restricts the variable \(z_{ji}\) to binary values.

5. Model testing

The model is tested with the actual data obtained from a company that considered reconfiguring its distribution network. The network consists of 1 plant located at Terre Haute, Indiana, and 21 warehouses serving 281 customers located throughout the United States and Canada. The warehouse and customer locations were given by city and zip code. The cities hosting the warehouses, 18 in the United States and 3 in Canada, are listed in Table 1. The company is one of the largest North American manufacturers and distributors of plastic film used (i) for packaging such items as bakery goods, and cigarettes, and (ii) for labeling soft drink bottles and cans. To honor the wish of the company to remain anonymous, the company will be referred to as the ‘firm’.

The firm provided customer demands and warehouse capacities and inventories over a 1-year period. Most of the warehouses were substantially underutilized under the current distribution network. According to the warehouse-utilization ratio (the total number of pallets at the warehouse divided by the theoretical capacity of the warehouse during the given period) used by the firm, all warehouses have less than 50% of the average annual utilization ratio for the last few years. Since the firm maintains the number of pallets sufficient to accommodate unexpected demand surge, some pallets may remain empty throughout the year. However, the product that the firm deals with does not have dramatic seasonal peaks and valleys, the firm seldom needs the number of pallets to exceed the average annual demand.

Table 1

<table>
<thead>
<tr>
<th>i</th>
<th>City</th>
<th>State</th>
<th>Zip code</th>
<th>Intangible benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Terre Haute</td>
<td>IN</td>
<td>47408</td>
<td>68</td>
</tr>
<tr>
<td>2</td>
<td>Greenville</td>
<td>TN</td>
<td>37744</td>
<td>62</td>
</tr>
<tr>
<td>3</td>
<td>Memphis</td>
<td>TN</td>
<td>38118</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>Richmond</td>
<td>VA</td>
<td>23231</td>
<td>63</td>
</tr>
<tr>
<td>5</td>
<td>Paterson</td>
<td>NJ</td>
<td>07501</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>Orlando</td>
<td>FL</td>
<td>32824</td>
<td>70</td>
</tr>
<tr>
<td>7</td>
<td>Bensenville</td>
<td>IL</td>
<td>60106</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>Indianapolis</td>
<td>IN</td>
<td>46242</td>
<td>65</td>
</tr>
<tr>
<td>9</td>
<td>Little Rock</td>
<td>AR</td>
<td>72114</td>
<td>90</td>
</tr>
<tr>
<td>10</td>
<td>Minneapolis</td>
<td>MN</td>
<td>55414</td>
<td>58</td>
</tr>
<tr>
<td>11</td>
<td>Winnipeg</td>
<td>MB</td>
<td>R2H 0Z7</td>
<td>46</td>
</tr>
<tr>
<td>12</td>
<td>Montreal</td>
<td>QC</td>
<td>H4Y 1E7</td>
<td>38</td>
</tr>
<tr>
<td>13</td>
<td>San Leandro</td>
<td>CA</td>
<td>91710</td>
<td>40</td>
</tr>
<tr>
<td>14</td>
<td>Toronto</td>
<td>ON</td>
<td>M8Z 5P7</td>
<td>35</td>
</tr>
<tr>
<td>15</td>
<td>York</td>
<td>PA</td>
<td>17403</td>
<td>55</td>
</tr>
<tr>
<td>16</td>
<td>Dallas</td>
<td>TX</td>
<td>75228</td>
<td>75</td>
</tr>
<tr>
<td>17</td>
<td>Platteville</td>
<td>WI</td>
<td>53818</td>
<td>58</td>
</tr>
<tr>
<td>18</td>
<td>New Castle</td>
<td>DE</td>
<td>19720</td>
<td>46</td>
</tr>
<tr>
<td>19</td>
<td>Fond du Lac</td>
<td>WI</td>
<td>54935</td>
<td>57</td>
</tr>
<tr>
<td>20</td>
<td>Laredo</td>
<td>TX</td>
<td>78045</td>
<td>80</td>
</tr>
<tr>
<td>21</td>
<td>Atlanta</td>
<td>GA</td>
<td>30336</td>
<td>70</td>
</tr>
</tbody>
</table>
To obtain some idea as to how large the customers are and how far away they are located from the warehouses, the annual customer demand versus access time from the closest warehouse is plotted in Fig. 2. Therein, a dot represents a customer, and demand is in cwt units (1 cwt = 100 lbs). Access times between warehouses and customers, identified by their respective zip codes, were obtained on the highway transportation network of the United States and Canada using a Geographic Information System at the University of Louisville. As Fig. 2 shows, the majority of the customers have demands less than 70,000 cwt, and can be accessed within 10 hours delivery time from one of the warehouses (many dots overlap near the origin). However, (i) there are five large customers with demands ranging from 80,000 to 250,000 cwt, and (ii) there are eight relatively small customers with access times of more than 10 hours and as high as 23 hours. Some cost data were provided by the company but the bulk of them were estimated by the authors using several indices, such as cost of living index (e.g. Savageau and D’Agostino, 2000). The warehouse capacities were assumed equal to 250,000 cwt. The cost of living index was used together with the quality of life index to estimate the intangible benefits of the warehouses at their various locations (last column of Table 1).

In order to set the desirability ranges, we first evaluated the three objectives for the current distribution network (status quo) to gain a sense of the current performance, and what the reasonable expectations could be. The status quo values were $68.05 million, 98.05% and 59.78 for the cost, demand coverage, and intangible benefits, respectively. The criteria preference-ranges were then set accordingly. For example, for the cost criterion, the most ideal situation would be a decrease in total costs by at least 15%, thus $t^{+}_{p1} = 0.85 \times 68.05 = 58.8$. Similarly, it was thought that a 5% cost decrease was desirable ($t^{+}_{p2} = 64.6$), a 2% decrease tolerable ($t^{+}_{p3} = 66.6$), while any decrease below 2% was undesirable (making it not worthy of pursuing network restructuring). On the other hand, a cost increase from its current level ($t^{+}_{p4} = 68.05$) is highly undesirable and a cost increase of 3% or more ($t^{+}_{p5} = 70.1$) was considered unacceptable. Given the ranges for each criterion, the associated incremental weights are derived using the LPP algorithm (e.g. Messac et al., 1996). The desirability ranges along with the associated weights (in italic) are shown in Table 2.

The resulting single objective mixed-integer linear programming model, displayed in (5)–(24), has 6381 continuous variables, 441 binary variables and 385 constraints. It was solved by the commercial software package, LINGO (e.g. LINDO, 2001). It took 38 minutes to solve on a Pentium III PC. The use of a speedier Pentium IV PC would have solved the problem even more quickly. Since both LINGO software and Pentium III PC are readily available to the company executive, the proposed model can be applied to many practical settings dealing with warehouse restructuring.

As Table 3 shows, the baseline model ($\tau = 10h$) suggests to (a) phase-out 8 warehouses, (b) relocate 4 warehouses and consolidate them with 3 other existing warehouses, and (c) maintain 6 warehouses at their current locations without any changes. In particular, the following warehouse consolidations are suggested by the model: the warehouse at Terre Haute, IN, is moved 65 miles to be merged with the warehouse at Indianapolis, IN; the warehouses at Platteville and Fond du Lac, WI, are moved 159 and 152 miles, respectively, to be merged with the warehouse at Benseville, IL; and the warehouse at New Castle, DE, is moved 73 miles to be merged with the warehouse at York, PA. In all cases, the relocation distance is less than
160 miles, so that relocation costs will be relatively low, and disruption of supply chain operations during transition less severe. In the map of Fig. 3, warehouses are indicated with numbers ranging from 1 to 21 and customers with dots. The open warehouses in the downsized distribution network are circled. Arrows originating at the warehouses to be relocated and directed toward the warehouses to be consolidated signify relocation for consolidation. We note that the majority of warehouses suggested for phase-out or relocation are either underutilized, or have low intangible benefits and high costs. It is also interesting to note that most of the retained warehouses, including the consolidated ones, are either at the center, or in the vicinity of the center of concentrated demand locations.

Fig. 3 also displays with different shades the geographical regions served by open warehouses. The consolidated warehouse (#15) in York, PA, will serve customers located in Eastern Pennsylvania, New Jersey, Delaware, New York and New Jersey.
England States. The Greenville-based warehouse (#2) will serve customers located in Eastern Tennessee, Eastern Kentucky, Virginia, Maryland, West Virginia, North Carolina and South Carolina. The Atlanta-based warehouse (#21) will serve customers located in Georgia, Florida, Alabama, Mississippi, Western Louisiana and all but Eastern Tennessee. The consolidated warehouse (#8) at Indianapolis, IN, will serve customers located in Missouri, Southern Illinois, all but Eastern Kentucky, Indiana, Ohio, Michigan, Northern Pennsylvania, Southeastern Ontario and Quebec, Canada. The Little Rock-based warehouse (#9) will serve customers located in Arkansas, Western Louisiana, Northern Texas, Oklahoma, New Mexico, Colorado and Southern Oklahoma. The Laredo-based warehouse (#20) will serve all customers located in Texas, except Northern Texas. The San Leandro-based warehouse (#13) will serve customers located in California, Nevada and Arizona. The consolidated warehouse (#7) at Bensenville, IL, will serve customers located in the highly concentrated demand regions of Northern Illinois and Wisconsin, as well as isolated customers in the Midwestern and Northwestern United States. Finally, the Winnipeg-based warehouse (#11) will serve a high-volume customer in the same city, which is using 99% of its capacity, and a few small customers scattered over Midwestern and Western Canada from Ontario to the West Coast.

Table 4 shows the improvements achieved in the three measures of performance for the restructured network, and the base line scenario (τ = 10 hours) over the current network configuration, consisting of all 21 warehouses and assuming optimal demand allocation to customers. The cost shifts favorably from the intersection of the undesirable and highly undesirable ranges, $68.05 million, to $61.457 million, which is in the desirable range. The demand covered within 10 hours access time worsens from 98.05% to 97%, but it remains in the ideal range. This can be explained by the fact that the DM considers any value greater than or equal to 97% as ideal. By reducing the covered demand to 97%, however, better values for the other objectives (better tradeoff) can be achieved. Finally, the intangible benefits improved by shifting

Fig. 3. Optimal distribution network.
from the undesirable range (59.78) to the desirable range (66.03).

**Table 3** illustrates the sensitivity of the optimal network configuration to changes in the parameter \( \tau \). As access time increases from \( \tau = 10 \) to 11 hours, the optimal configuration remains unchanged. Further increase to \( \tau = 12 \) hours changes only the mix of consolidated warehouses. The total costs are reduced to $60.88 million mainly due to decrease in inbound shipping costs (the warehouse at Terre Haute is in the same city as the plant), while 97% of demand (ideal) is served within 12 hours of access time. On the other hand, in order to maintain a desirable performance in the coverage criterion, as \( \tau \) decreases from 10 to 7 hours, two fewer warehouses are phased-out and only two (instead of three) warehouses are consolidated. In the event of more stringent accessibility requirements (\( \tau = 7 \) hours), the model finds a solution that worsens all criteria but in a balanced way, i.e., all objective values are in the desirable range. As compared to the base line scenario, the cost increases to $63.92 million, the covered demand within 7 hours decreases to 93.65%, and the intangible benefits decrease by one point to 65.

### 6. Summary and conclusions

The reconfiguration of a distribution network entails the consideration of all pertinent factors. These include: relocation/consolidation costs, inbound and outbound transportation costs, warehousing costs, relocation costs, customer delivery time, and intangible factors, such as labor quality, labor-management relations and tax incentives. In this paper, these factors are combined in a multi-criteria model that can aid the management of a company in reconfiguring its distribution network as part of downsizing (or "right-sizing") decisions. In addition to the new distribution network of plants and warehouses, the model determines the demand allocations of plants to warehouses and warehouses to customers. The model was applied to a company that wanted to restructure its distribution network over the USA and Canada. For potential implementation, the model results were presented to the company executive and management team. Since the model implementation entails several years of organizational re-structuring and business reengineering, the firm in this case study has not fully implemented the detailed phase-out plan recommended by the model builders at this stage. However, the company executive and management team acknowledged that solutions based on the proposed model clarified the potential cost–benefit accrued from the warehouse restructuring decision.

The recently developed LPP approach was used to formulate the multi-criteria model. Under the LPP framework, the decision maker expresses his/her preference for objective function values in terms of ranges of differing degrees of desirability. Specifically, the DM provides numerical values that define each of these ranges for each criterion. Thus, LPP removes the necessity to choose weights—which is undoubtedly a difficult task—and allows decision makers to reflect their preferences in an explicit and flexible manner.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Ideal</th>
<th>Desirable</th>
<th>Tolerable</th>
<th>Undesirable</th>
<th>Highly undesirable</th>
<th>Unacceptable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost in US$ 10^6</td>
<td>0</td>
<td>58.8</td>
<td>64.6</td>
<td>66.6</td>
<td>68.05</td>
<td>70.1</td>
</tr>
<tr>
<td>Customer coverage</td>
<td>100</td>
<td>97</td>
<td>90</td>
<td>80</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>Intangible benefits</td>
<td>100</td>
<td>80</td>
<td>65</td>
<td>55</td>
<td>45</td>
<td>40</td>
</tr>
</tbody>
</table>

- Original network; ○ optimized network.
Despite numerous merits, the proposed model points to a number of directions for future work:

1. The model can be extended to include multiple products, risk and uncertainty of costs and other factors, as well as customer demand uncertainty over the time horizon.
2. Regarding the LPP approach, sensitivity analysis tools of the user’s specified preference ranges should be developed. Initial work in the non-linear Physical Programming domain shows interesting and promising results (e.g. Tappeta et al., 2000). They indicate that the obtained optimal solutions are not unduly sensitive to the range settings.
3. More efficient solution techniques for solving the resulting large mixed-integer linear programming problem should be developed, especially when the above extensions are implemented and additional runs are needed for performing sensitivity analyses.

Acknowledgements

The authors are grateful to the management team of the anonymous firm for providing us with the data and their insight into the distribution network restructuring problem and to the two anonymous referees who provided valuable suggestions for improving this paper. Support for Dr. Achille Messac under the NSF grant number DMI 0196243 is hereby acknowledged.

Appendix

The following are the three most important properties of class functions:

(I) A class function is nonnegative, continuous, piecewise linear, and convex,

(II) the value of a class function, \( z_p \), at a given target level (say, \( t_{p3}^+ \)) is the same for any class-type, and

(III) the magnitude of the class function’s vertical excursion across any range must satisfy the One versus Others (OVO) rule.

The first property is evident. The second property means that, as one travels across a given range type (say, undesirable range-4), the change in the class function will always be of the same magnitude, \( \tilde{z}^4 \) (see Fig. 1), regardless of the criterion in question. This behavior of the class function at the target levels is a critical factor that makes each range type have the same numerical penalty value for different criteria. This same behavior also has a normalizing effect and results in favorable numerical conditioning properties. For example, consider the behavior of two distinct criteria over the undesirable range. The first criterion values vary between 5000 and 12,000, while the second vary between 1.9 and 4.2. Since both of those ranges are undesirable, the respective class functions will change over that range by the same amount, \( \tilde{z}^4 \). The third property needs some clarification. The OVO rule entails the following inter-criteria preference for each criterion, \( g_p \): If two options are considered, viz.,

Option 1: “Full improvement of \( g_p \) across a given range (say, range-3)”; and

Option 2: “Full reduction of all the other criteria across the next better range (i.e., range-2)”; then option 1 shall be preferred over option 2. This is to say that the worst candidate entails a higher penalty.

A comment on the ideal range: we wish to explicitly emphasize the point that, according to the definition of the ideal range, any two points of the ideal range are of equal value/desirability. Consider the case of Class 1S. The class function will seek to minimize its criterion only until the target value \( t_{p1}^+ \) is reached. Below that point, Class 1S expresses explicit indifference. If a smaller value of the criterion is always better, the ideal range does not apply. In this case, the DM should set \( t_{p1}^+ \) to a value outside of the feasible space in order to exclude solutions in the ideal range. By being mindful of the above comment, the DM will easily preclude the possibility of obtaining (incorrect) dominated solutions. A similar discussion would apply to the cases of Classes 2S and 4S. For more details on the class function properties and their rationale, see Messac (1996).
References


