A MILP facility location model with distance value adjustments for demand fulfillment using Google Maps

Kevin Palomino *, David Garcia ** and Carmen Berdugo***

* Assistant Professor, Department of Industrial Engineering, Universidad del Norte, Colombia.
** Independent Researcher, consultant, and trainer, Industrial Engineer, Universidad del Atlántico, Colombia.
*** Researcher, Department of Industrial Engineering, Universidad del Norte, Colombia.
* Corresponding Author: krpalomino@uninorte.edu.co

ABSTRACT

In this article, a facility location model was designed to support logistics operations, considering service distance limitations for demand fulfillment and a list of candidate locations within a supply chain. Consequently, an allocation model was designed using Mixed-Integer Linear Programming (MILP), in which a finite number of demand nodes could be satisfied by a set of supply nodes, considering not only the costs related to these locations, but also restrictions aimed at improving the level of service based on distance. Besides, an integrated solution scheme was proposed that includes a macro in VBA language that calculates the distance between nodes using the web mapping service developed by Google Maps and solving the model through a branch and cut algorithm. Subsequently, a case study was executed, where the supply operation of an important Colombian retail company is analyzed. The results reflected positive effects not only on costs, but also on the prioritization of average distance traveled and on the satisfaction of store demand by distribution centers. Thus, the conditions in which the implementation of this model provides strategic benefits were verified, functioning as a tool to support decision making.

Keywords: Facility location; Logistics engineering; MILP; Supply chain management.

INTRODUCTION

The main objective of this research is to design a model of facility location with distance limitations, capable of providing real and effective solutions for the design of supply networks. This paper addresses three specific research questions: (1) How can the current facility location models be improved? (2) How can the tendency to look locations on a continuous surface without considering the reality and context of why a location has been selected be prevented? (3) Is minimizing total costs generally a primary goal in making facility location decisions?

The reason this topic is important is because of its ability to manage complex logistics networks, and it is considered crucial to the success or failure of a business. Hence, the correct design of the network where tangible and intangible resources flow within a supply chain has become necessary to create added value and competitiveness.

Besides, it is imperative to mention that one of the main decisions faced by logistics management is in the optimization of the physical flow of goods, or more specifically, in the selection of facilities. The Council of Supply Chain Management Professionals (CSCMP) defines the term 'facilities' as an installation or contrivance that facilitates something; a place for doing something: Commercial or institutional buildings, including offices, plants, and warehouses (CSCMP, 2013). Also, facility location plays a vital role in supply chain design, as it impacts their long-term logistics costs (Yang et al. 2019). A facility location problem consists of deciding where to locate one or several facilities in order to serve a set of demand points. Often, the goal is to minimize the total cost that includes establishing the facilities and supplying the demand (Kınay et al. 2019). Thus, facility location is one of the very first and prominent strategic decisions that profoundly affect operational decisions in any organization (Farahani et al. 2019).

Based on the above, the research on facilities location evolved towards interdisciplinary where not only complex models based on mathematical programming are implemented, but also, on complex algorithms and heuristics, this is explored further in section 2, literature review. Moreover, the model background embraces methodologies such as the Weber method or the calculation of the center of gravity, which aims to find a location within a continuous surface. Methods that still lack the accuracy to meet the current needs of companies. Since one does not live in a Euclidean space and there is no certainty that a precise location thrown by such methods is the most indicated or desirable. Today, other aspects must be considered, such as the level of service, the road and rail network, or the possibility of acquisition that a company may have on a property, for example. These traditional methods ignore these aspects. So, in practice, there is a need to design more accurate models on a large scale, because these models should not only look at a location on a continuous surface, it should consider the reality and the context on why a location is selected.

The possibility of adjusting the capacity of operating facilities over the planning horizon is relevant when sizing decisions are reversible in the medium term. In this context, it is important to understand the trade-off between the costs incurred by capacity expansion and contraction and the level of service provided to the customers. For example, it may be economically attractive to invest in increasing capacity temporarily to guarantee shorter delays in demand fulfillment (Correia et al. 2017). In the tactical level of logistic planning, level of service could be defined as a Cycle Service Level, it looks at the probability of stockout during a replenishment time, or as an item fill rate, the percent of the items ordered that come in. However, on strategic design, distance to the customer can be a proxy for archiving those goals. Furthermore, this is a good way that it can get at in measuring the opportunity that any firm can meet these tactical service objectives.

Therefore, as the main contribution, a solution scheme is proposed that comprises a Mixed-Integer Linear Programming (MILP) model solved through a branching and cutting algorithm. In parallel, the model involves a list of candidate installations and works with distance values taken from the web mapping service developed by Google Maps. Finally, for its validation, a case study was carried out with a Colombian retail company to select distribution centers (DC) to supply a set of stores nationwide.

These methods have been proposed considering that instead of having a continuous space where any type of location can be obtained, the model works with a set of candidate locations. This is based on the fact that it does not make much sense to obtain locations where a facility cannot be located or where a property can not be acquired. Just to mention an example, a traditional model could optimally result in one of the facilities being located on a lake, or it does not consider whether the company has the possibility of quickly acquiring a specific property in a specified location. Then, these candidate locations provide input to the model, so it is indicated that they should be the result of a previous feasibility study by the company that could involve a study of investment alternatives to explore the conditions, availability, and cost of such facilities.

Besides, the results of the validation carried out presented positive benefits concerning the implementation of the proposed model to solve problems of selection of logistics facilities, proposing a methodology that considers

viable location options for the company and that is a product of its own needs. Similarly, not only was it possible to establish in this allocation model a prioritization of the costs associated with this activity, but also restrictions were considered aimed at improving the level of service using the distance between supply and demand nodes, together with percentages of this demand satisfied within certain distance limits. Also, synergic work was achieved with real data from Google's geographic information system (GIS) to collect distance information.

This article is organized as follows: The related literature is reviewed in Section 2; Section 3 describes the methodology used in this model; Section 4 explains each of the stages of model design; Section 5 concludes this paper and gives some recommendations for future studies.

RELATED WORK

Facility location problems received attention from many perspectives in the scientific literature and remained relevant nowadays. Some researchers study these problems from different design methods. For example, Khosravi and Jokar (2017) model adheres to gravity rule for the domestic facility and hub location in a transportation network where the number of facilities is unknown, and the goal is to minimize the transportation cost, lost demand penalty and facility set up costs. From the robust optimization (RO), Jakubovskis (2017) presented a modeling approach to develop insights into strategic capacity planning and resource acquisition decisions, including the facility location problem. Other authors rely on the Weber facility location problem, this is the case of Drezner, Drezner, and Schöbel (2018) who propose the Weber obnoxious facility location problem whose, as in the classic Weber location problem, the objective is to minimize the weighted sum of distances between the facility and demand points. However, the facility location is required to be at least a given distance from demand points because it is "obnoxious" to them. Another example is the study of Du, Zhou, and Leus (2020), where a two-stage robust model is designed based on the p-center location model for reliable facility location when some facilities can be disrupted. Thus, a reliable network is constructed in a "proactive" planning phase, and when a facility is disrupted, its original clients can be reallocated to another available facility in a "reactive" phase.

In that context, linear programming became widely used in the construction of this type of model. Boujelben, Gicquel and Minoux (2016) studied a multiperiod facility location problem, including some realistic constraints to take into account vehicle routing from distribution centers to customers while keeping a manageable size of the optimization problem which is formulated as a mixed-integer linear program and solved using a commercial solver. Also, Correia and Melo (2016) addressed an extension of the classical multiperiod facility location problem in which customers are sensitive to delivery lead times. They propose two mixed-integer linear programming formulations to re-design the network at minimum cost. Similarly, Orjuela-Castro, Sanabria-Coronado, and Peralta-Lozano (2017) propose a mixed linear programming model for the localization of collection centers and companies processing perishable foods in mountainous regions, based on a multiproduct and multiechelon transport system. Teye, Bell, and Bliemer (2017) analyze a problem that comprises a mode choice problem embedded within a facility location problem. They employ the method of entropy maximization to combine a logit mode choice model with a facility location model, leading to a nonlinear mixed-integer programming model. Another study worth mentioning is Emirhüseyinoğlu and Ekici (2019), who formulated a mixed-integer mathematical model a problem related to multiperiod facility location decisions of a retailer which procures the products from multiple suppliers under an incremental quantity discount scheme and in turn satisfies an exogenous demand. On the other hand, Golpîra (2020) proposes an original Mixed Integer Linear Programming to integrate Vendor Managed Inventory strategy into the general multiproject multiresource multisupplier Construction Supply Chain (CSC) network design and facility location problems at a minimum cost.

Along the same lines, there are also contributions from nonlinear programming. In the work of Hajipour, Fattahi, Tavana, and Caprio (2016), a multiobjective multilayer facility location-allocation (MLFLA) model with congested

facilities using classical queuing systems is proposed to determine the optimal number of facilities and the service allocation at each layer. Similarly, Qi, Xia, Zhang, and Miao (2017) present a bilevel, nonlinear, integer programming model for the competitive facility location problem with foresight. The developed model's objective is to maximize the leader's market share while also taking into consideration the follower's response. Coniglio, Fliege, and Walton (2017) consider a facility location problem to determine the optimal number and location of warehouses and which items are to be stored in each of them. They propose a nonlinear mixed-integer programming formulation for the problem.

Additionally, Yu, Haskell, and Liu (2017) consider a reliable uncapacitated facility location problem (RUFL) with random facility disruptions. They develop risk-averse optimization formulations to compute resilient location and customer assignment solutions. Likewise, Karatas (2017) presents a multiobjective facility location problem, which includes facilities with gradual covering decay, cooperative demand coverage, and variable coverage performance. The location problem is modeled as a multiobjective integer nonlinear program (INLP). Also, Tadros, Galal, Ghazy, and ElSayed (2018) addressed the multiobjective two-echelon capacitated facility location problem under covering distance restrictions. It aims to minimize total cost and to maximize demand coverage for a hierarchical facility network while considering covering distance.

Researches also focused on goal programming. Cedolin, Göker, Dogu, and Esra Albayrak (2018) solved a facility location selection problem in a manufacturing company. For this purpose, the facility location selection problem was solved by applying fuzzy data envelopment analysis (Fuzzy DEA) and fuzzy goal programming (Fuzzy GP) methods. Karatas and Yakıcı (2018) presented a novel methodology for solving multiobjective facility location problems. The proposed approach is mainly based on a combination of the branch and bound and iterative goal programming techniques. Besides, research using game theory was also conducted in the more recent literature, such as the study of Rohaninejad, Navidi, Nouri, and Kamranrad (2017) that deal with cooperative competition in facility location problems in which potential players (investors) compete in acquiring suitable sites and clients.

Similarly, Nasiri, Mahmoodian, Rahbari, and Farahmand (2018) introduced the capacitated competitive facility location problem, which is a typical optimization problem that is also associated with the game theory. In essence, the problem is composed of two competitors who seek to attract customers by establishing new facilities and maximizing their profit. It should also be mentioned the investigation of Mei, Li, Ye, and Zhang (2019) who explore two types of facility location games: the facility location game where each agent wants to minimize their distance from the facility, and the obnoxious facility game where each agent prefers to be as far away from the facility as possible.

Other researches focused on complex algorithms and heuristics. For example, Guo, Cheng, and Wang (2017) considered the two-stage capacitated facility location problem (TSCFLP) in which products manufactured in plants are delivered to customers via storage depots. They proposed a hybrid evolutionary algorithm framework with machine learning fitness approximation for delivering better solutions in a reasonable amount of computational time. Manthey and Tijink (2018) analyzed a simple local search heuristic for the facility location problem using the notion of perturbation resilience. Besides, Lin et al. (2018) considered a facility location-allocation problem of end-of-life vehicle recovery networks and established a mathematical model based on an artificial bee colony to solve the problem. The model achieved the minimization of cost for deciding optimal locations of end-of-life vehicles recovery network. In another study, Chalupa and Nielsen (2019) proposed a new Monte Carlo hybrid local search algorithm (Hyb-LS) for solving the uncapacitated facility location problem. Hyb-LS is based on repeated sampling using two local search strategies based on best improvement and randomized neighborhood search. Afify et al. (2019) propose an evolutionary learning technique to solve two research problems near-optimally: Reliable p-Median Problem and Reliable Uncapacitated Facility Location Problem considering heterogeneous facility failure probabilities, one layer of backup and limited facility fortification budget.

In turn, Biajoli, Chaves, and Lorena (2019) presented a new metaheuristic approach for the two-stage capacitated facility location problem (TSCFLP), in which the purpose is to minimize the operation costs of the underlying two-stage transportation system, satisfying demand and capacity constraints. Finally, Sauvey, Melo, and Correia (2019) investigated a recently introduced extension of the multiperiod facility location problem that considers service-differentiated customer segments: some customers require their demands to be met on time, whereas other customers accept delayed deliveries as long as lateness does not exceed a pre-specified threshold. They propose four heuristics that construct initial solutions to this problem and subsequently explore their neighborhoods via different local improvement mechanisms.

In summary, the literature on facility location modeling has been growing, and the **increasing** interest in these problems is motivated by the recognition of the need to consider more criteria to achieve solutions that are closer to reality. Compared to the proposed model, the main advantages of these models mentioned are related to providing a quantitative basis for a systematic analysis of facility location problems and restricting the decision space. Also, these models divide the decision process into multiple steps, minimize subjectivity for the decision, and use the information provided in the best possible way. On the other hand, the most prominent disadvantages are associated with the need to require careful consideration of the problem and correctly lead to the elimination of unfeasible locations or the accentuation of desirable alternatives. In nowadays' world, thinking about objective functions other than economic functions is becoming a necessity, and the integration of location decisions with other network design decisions, such as routing, resource utilization, and customer responsiveness, can be of significant development.

METHOD

This section presents the main guidelines of the proposed model. The problem of designing a facility location model was analyzed by considering a list of candidate facilities for selection, in order to avoid falling into the unrealistic scenario of searching for locations within a continuous surface by traditional methods such as the center of gravity or the Weber method. A network can be defined as a set of nodes and arcs, where a node is a point within the network, and an arc is a junction between two nodes (Cadarso et al. 2017). Considering the above, this model was designed as a selection and assignment model, in which a finite number of demand nodes could be satisfied by several supply nodes, considering costs and service level. In other words, we sought to find the optimal number of facilities that would satisfy the demand for other facilities.

Although these concepts may seem simple, solving an optimization problem of this type involving real data can be a complicated task. For example, if a company has 500 stores dispersed in a given area and has five candidate locations for its distribution centers, it would be necessary to analyze 2,500 possible supply combinations that may involve distance variables or demand flows, not to mention the difficulty of solving the problem analytically. Therefore, an integrated solution scheme was proposed, focusing on the design of a Mixed-Linear Programming Model to achieve the proposed objectives. Regarding managerial implications, strategic decisions are primarily concerned with the development of operational plans and the establishment of policies and management of resources to meet current and future operational needs and customer requirements, in a manner consistent with organizational goals. The proposed model allows management to define facility service levels and evaluate the trade-off with total cost. For example, facilities that are strategically located or have high capacity are expected to have higher service levels. In brief, facility location is a strategic problem that is part of the planning process for decision-making in supply chain design.

Finally, since most of the real problems of this type are not susceptible to be solved by traditional methods or spreadsheets, the model was solved with SAS, a computer programming language created by SAS Studios. This software is recognized as one of the leaders in statistical analysis and as a powerful tool for database systems. The main reason why SAS was selected among other software packages was that it includes several optimization tools for a wide range of operations research problems (Semeida 2014; Catal et al. 2015; Ding et al. 2016; Pusponegoro et

al. 2017; Rachmawati et al. 2017; Wang et al. 2017). Besides, a macro was designed in VBA language to calculate the distance between nodes using the Google Maps web mapping system. This macro was based on the publicly available access to cartography and actual data of the road network for a specific study problem, then exporting these data remotely to Microsoft Excel. All this was done through the Internet, using a standard browser.

Model Formulation

First, in formulating the model, indices i and j were defined, denoting a specific supply node and a demand node, respectively. The notation x_{ij} represents the flow from node i to the node j. Also, there are a total of n supply nodes and a total of m demand nodes. The objective function (z) is presented in Equation (1), in order to minimize the total costs. It was considered a fixed cost (C_i) that is incurred when an O_i facility or supply node is used or 'opened'. Similarly, a variable cost related to the transportation cost (c_{ij}) from the node i to the node j. This value is multiplied by the x_{ij} flow.

Note that O_i is a binary variable. This means that it can only take two values, 0 or 1. Thus, this variable was added to the target function and is multiplied by the fixed cost of opening an i-node in a given time horizon, so C_i is the cost of opening this facility for that period. Also, a capacity constraint defined in Equation (2) was formulated in order to ensure that the total available capacity within the whole set of those supply nodes (N_i) is not exceeded. Similarly, Equation (3) denotes the demand constraint. To guarantee that all flows entering a given node j must exceed or equal the demand at each of them (D_j). In addition, two constraints limiting the number of open facilities were introduced into the model: Equation (4) sets a minimum value, F_{min}, and Equation (5) sets a maximum value, F_{max}. These values are the number of installations that could be opened.

In addition, a tying restriction was necessary, because it is not possible to supply from a supply node unless it is open. After all, the objective function will force its closure. In equation (6), B_{ij} represents a large number, and the binary variable O_i can take the values 0 or 1. It is important to mention that it is not recommended to assign a very large value to B_{ij}, because it could slow down the solution time. For this reason, it was proposed that it be equal to the sum of all the demand since the flow in each arc can never be higher than the total demand required. Beyond that, when deciding whether to open several facilities, it is necessary to consider the advantages and disadvantages of a location model. In any type of supply chain design, a balance must be struck between the interests of the competition, and this is mainly achieved with competitiveness in costs and services. There are several types of costs that a company may incur, and service is related to customer satisfaction. Companies want to make sure that their customers are satisfied, and when they order something, they get it. In turn, companies want to achieve growth in their sales, expand their market share, among other proposed challenges.

Thus, the models of location of facilities are included within the strategic level of decision making that considers the uncertainty of the market conditions foreseen for the coming years (Chopra et al. 2016). Therefore, the aim is to find a scenario that improves profitability, which means being attentive to costs and increasing customer satisfaction and service. The term "service level" could mean different things to different people and companies. At the tactical level of logistics planning, the service level considers the probability of stock-outs during a replenishment time. However, because the research focused on strategic design, the model used distance to the customer to achieve these service objectives. A shorter distance between supply and demand facilities is related to the role of coordinating supply and demand imbalances, reducing transportation costs, and adding value to products through rapid service, among other benefits. The model includes two related constraints, Equation (7) and Equation (8).

Equation (7) is a restriction for the maximum allowed mean distance. Each x_{ij} is multiplied by the distance d_{ij}, which represents the distance to node j from the node i by determining a weighted distance, this value is divided by the total demand (D_j) to calculate this percentage. The constraint adds up these quantities for all x_{ij} values and gets

a weighted average distance that must be less than or equal to a value ϕ. The value ϕ is the maximum average distance and was considered an input value. It is imperative to mention that this value ϕ is not determined for all nodes i or for all nodes j. It is only a distance limitation in the model. Thus, the maximum average distance is established for all flows in the network, and customers or regions with higher demand are being favored. It should be stressed that the model needs a distance matrix as input data for this constraint. This is achieved by running the macro in VBA explained in section 3.2.

The next constraint is the minimum allowable demand within some specified distance value, Equation (8). The value is above a certain amount φ. So, the right-hand side of this constraint is an input value. The left-hand side of this constraint includes the flow through x_{ij}, and multiplying that by a new input data, which is a constant a_{ij}. This value is equal to 1 if a demand node j served by a source node i is within a minimum allowable demand value φ. It is 0 otherwise. The values of a_{ij} are included in a matrix with all possible combinations. For example, if a model is designed for a logistics network with distribution centers and stores, the distribution centers are the source nodes, and the stores are the demand nodes, and a value φ of 50 miles is considered. If a distribution center i serve a store j, then a_{ij} will have the value of 1 if it is within 50 miles; otherwise, it is 0. Then, if the model is serving that destination from the selected distribution center, it is either within 50 miles or not. The constraint only counts the ones that are within 50 miles, and it divides that value by the total demand. Then, it is taking the percentage of the total demand that is within 50 miles. Finally, Equation (9) presents a nonnegative restriction for flows between nodes. The complete model is presented as follows:

$$z = \sum_{i=1}^{n}\sum_{j=1}^{m} c_{ij}x_{ij} + \sum_{i=1}^{n} C_i O_i \tag{1}$$

$$\sum_{j=1}^{m} x_{ij} \leq N_i \quad \forall i \in S \tag{2}$$

$$\sum_{i=1}^{n} x_{ij} \geq D_j \quad \forall j \in D \tag{3}$$

$$\sum_{i=1}^{n} O_i \geq F_{min} \tag{4}$$

$$\sum_{i=1}^{n} O_i \leq F_{max} \tag{5}$$

$$x_{ij} - B_{ij}O_i \leq 0 \quad \forall ji \tag{6}$$

$$\sum_{ij} \left(\frac{d_{ij}x_{ij}}{\sum_j D_j}\right) \leq \phi \tag{7}$$

$$\sum_{ij}\left(\frac{a_{ij}x_{ij}}{\sum_j D_j}\right) \geq \varphi \tag{8}$$

$$x_{ij} \geq 0 \quad \forall ij \tag{9}$$

VBA Macro

One of the necessary model inputs corresponds to the distances between the possible combinations of supply nodes and demand nodes. However, the availability and collection time of these data and the price of adequate mapping can be a problem for such a model. Besides, it requires a network structure defined in the maps and associated route optimization algorithms, a task that is not accessible by traditional methods. However, Google Maps services allow overcoming these limitations by providing Internet access to mapping, road networks, and actual data associated with traffic restrictions (Branson et al. 2018; Google LLC 2017b; Yang et al. 2016). Accordingly, a macro in VBA (Visual Basic for Applications) language was designed that integrates this geographic information system (GIS) based system with this web tool. Thus, it is possible to collect real distance data from the road network for any case study. The macro recorded this information in a Microsoft Excel spreadsheet.

Maps and navigation systems have many complex algorithms and heuristics. In this case, the Distance Matrix API is a service that provides travel distance and time for a matrix of origins and destinations, based on the recommended route between start and endpoints. Anyone can access the Google Maps API through an HTTP interface, with requests built through a URL (Uniform Resource Locator) using sources and destinations, along with an API key. This key identifies an application for managing request quotas. The Google Maps APIs are free for a variety of use cases (Google LLC 2017a).

By using the real-time information shared by users and incorporating the past data, Google servers derive traffic information such as traffic along roads, estimated travel times between origins and destinations, the popularity of places and identification of extreme traffic conditions, etc. The algorithm or methodology of estimating traffic parameters is not revealed by Google and kept black-boxed to the public. However, the traffic information is public via Google Maps APIs (Kumarage 2018). Figure 1 shows the process of collecting traffic information from the Google application programming interface.
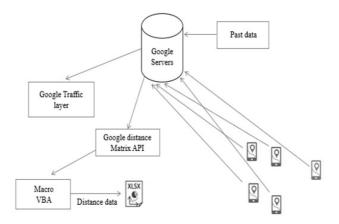

Figure 1. Collecting traffic data from Google Maps.

Therefore, the macro in VBA integrates this Google application to collect the distance information needed for the model and record this data in a spreadsheet. It is calculated through iterative API runs, providing actual travel distances between each pair of locations. By default, distances are calculated for the driving mode using the road network, but it should be noted that distances can be calculated that adhere to other modes of transport, with some modifications to the code. Similarly, Google allows to include restrictions on specific parameters to be avoided, for example, tolls, and motorways. Therefore, to design a model closer to reality, these types of restrictions were included. Likewise, one can specify the system of units to be used: metric or imperial units: the first returns distances in kilometers and meters, and the second, in miles and feet. In this case, the second option was chosen. The macro is detailed in Table 1.

Table 1. Macro in Visual Basic for Applications.

Linea	
1	Sub Algorithm_VBA()
2	Dim a, b, Str As String
3	Dim n As Integer
4	n = InputBox("Enter the number of arcs to calculate")
5	Dim lineS As Variant
6	On Error Resume Next
7	With CreateObject("WinHttp.WinHttpRequest.5.1")
8	a = CStr(ThisWorkbook.Worksheets(1).Range("A2"))
9	For j = 5 To (n+5)
10	b = ThisWorkbook.Worksheets(1).Range("a" & j)
11	.Open "GET","https://maps.googleapis.com/maps/api/distancematrix/json?units=metric&origins=" & a & "&destinations=" & b & "&key=YOUR_API_KEY_HERE", False
12	.Send
13	lineS = Split(.responsetext, vbLf)
14	lineS(8)= Trim(Replace(lineS(8), "text", ""))
15	lineS(8)= Trim(Replace(lineS(8), "mi", ""))
16	lineS(8) = Replace(Replace(Replace(lineS(8), Chr(34), ""), Chr(58), ""), ",", "")
17	lineS(8) = Trim(lineS(8))
18	ThisWorkbook.Worksheets(1).Range("b" & j)= lineS(8)
19	Application.Wait (Now + TimeValue("0:00:01"))
20	Next j
21	End With
22	End Sub

According to Table 1, the code lines are explained briefly. First, *Line* 1 corresponds to the instruction that executes the procedure in the active module. *Line* 2 declares two string variables that indicate the source and destination addresses or coordinates. Then, *Line* 3 declares the variable 'n' that will be the total number of arcs to be entered to calculate their distance. *Line* 4 is the instruction that allows the user to enter this number. *Line* 5 declares the variable *'LineS'* used to store the distance data. The instruction on *Line* 6 causes the execution to continue in case any geographic coordinate is incorrect or missing. 'With', in *Line* 7, executes a series of instructions that repeatedly refer to a single object. In this case, it creates a connection to the HTTP web server to send and receive data. Microsoft Windows HTTP Services (WinHTTP) provides developers with an application programming interface for sending requests over the HTTP protocol to other web servers.

In addition, *Line* 8 indicates that cell A2 contains the origin address (the user could change this cell for another one). This value is assigned to the variable 'a' by the function CStr, so that it is recognized as a string variable. String variables are used to store text. In *Line* 9, the repetitive cycle 'For' is used to cycle through the range of cells containing the destination addresses. *Line* 10 assigns the set of target addresses to a 'b' variable in each cycle. The 'Open' statement on *Line* 11 opens a connection to an HTTP resource. This statement has two parameters: 'GET' requests data from a specific resource, and the URL is the Google Maps web address or link. As for *Line* 12, the 'Send' statement sends a request to an HTTP server, connecting to the URL to calculate the distance between a source and a destination coordinate. *Line* 13 contains the 'response-text' property, which returns the Google Maps response code. The macro makes some adjustments to the code: it uses the 'Split' function to split the response code, separating it into many secondary lines as a one-dimensional matrix. It then uses the 'vblf' function, which generates a line break, thus separating the written text. These instructions are critical, as only the value of the distance is needed, and not the whole response code.

Regarding lines 14 to 17, the instruction *'lineS*(8)' is the line of the one-dimensional matrix created in the previous step that contains the value of the distance. The lines of code from 14 to 17 are in charge of perfecting the response generated, seeking to present only the value of the distance in the spreadsheet. The function 'TRIM' returns a text value with the initial and final spaces eliminated. In *Line* 18, the distance data is written to a set of cells in the spreadsheet (in this case in column B). The API can limit the requests in cases where one wants to get the distance between many addresses or coordinates. Therefore, the instruction in *Line* 19 serves to pause the execution until a specific time. *Lines* 20 and 21 end the execution of the 'For' cycle and the 'With' block. Finally, in *Line* 22, 'End Sub' ends the execution of the macro.

Solution Method

The mathematical formulation of the MILP model was written in SAS language to facilitate its solution (refer to Table A1 in the appendix). The executable model presented has no limitation in the number of nodes, since SAS has robust execution capabilities that allow handling large volumes of data efficiently and simultaneously. Similarly, this software includes a procedure called OPTMODEL that selects a solver subject to the type of problem to be solved (SAS Institute Inc. 2014).

Specifically, the Mixed-Integer Linear Programming (MILP) troubleshooter available from SAS runs a branch and cut algorithm. This algorithm was proposed by Land and Doig (1960) and is an effective method for solving this type of problem. The OPTMILP procedure implements a linear-programming-based branch-and-cut algorithm. This divide-and-conquer approach attempts to solve the original problem by solving linear programming relaxations of a sequence of smaller subproblems. The OPTMILP procedure also implements advanced techniques such as presolving, generating cutting planes, and applying primal heuristics to improve the efficiency of the overall algorithm. The OPTMILP procedure requires a mixed-integer linear program to be specified using a SAS data set

that adheres to the mathematical programming system (MPS) format, a widely accepted format in the optimization community (SAS Institute Inc. 2018).

The basic idea of branch-and-cut is breaking a problem into subproblems (sequences of LPs) that are easier to solve. Consider MILP:

$$J^* = \min_{(x,y)} c^T x + d^T y \; s.t. (x, y) \in X \tag{10}$$

where X is the set of feasible solutions,

$$X = \{(x, y) \epsilon \; \mathbb{R}_+^n \times \mathbb{Z}_+^p : Ax + By \geq b\} \tag{11}$$

Let $X = X_1 \cup X_2 \cup ... \cup X_k$ be a decomposition of the feasible solution set X into smaller sets X_k, and let $J^k = min\{c^T x + d^T y : (x, y) \epsilon \; X_k\}$ for $k = 1, ..., K$. Then $J^* = \max_k J^k$. The mathematical procedure for solving a particular mode is well-known; hence, it can be entirely omitted.

Thus, the OPTMODEL procedure provides a modeling environment adapted to build and solve optimization models. This makes the process of translating the symbolic formulation of an optimization model into OPTMODEL easy, as the modeling language imitates the symbolic algebra of the mathematical formulation as closely as possible. All this produces models that are more easily inspected for completeness and accuracy, more easily corrected, and more easily modified, either through structural changes or by replacing new data with old data (SAS Institute Inc. 2014).

CASE STUDY

In this case study, the proposed model was implemented to find the optimal configuration of a distribution network to meet the demand for products in several stores located in a territory. We studied a retail company located in Colombia that sells its products in 306 stores throughout the country. Currently, this company has five distribution centers (DC) located in the cities of Barranquilla (DC 1), Cartagena (DC 2), Medellín (DC 3), Bogotá (DC 4), and Cali (DC 5). Thus, the locations of these facilities were established as candidates in this case studying. It was also important to analyze and compare how their current distribution model differs from the optimal distribution model designed, in order to find the optimal location and number of distribution centers that will minimize costs and improve the level of service. In this particular case, only the demand requirements of the products within the food category are analyzed, as they represent the largest percentage of sales. Therefore, to include the consolidated demand for all products that fall under the food category, the demand volume was quantified in cubic meters (m^3). The capacity of the distribution centers was also quantified in this volume unit. Fixed costs corresponded to the costs of opening a given distribution center, and variable costs corresponded to transportation costs, set at USD 0.5 per mile.

Correspondingly, since the problem included 306 stores and five distribution centers, as shown in Figure 2, the macro designed to calculate the 1530 distance combinations was used. The macro was run with the exact coordinates of the stores to achieve a realistic solution. For the input data for the service level restrictions, the maximum allowable average distance was set at 150 miles, so the weighted average distance had to be less than or equal to that value. Similarly, the percentage of demand within 150 miles had to be a maximum of 85 percent. The case was settled at SAS.
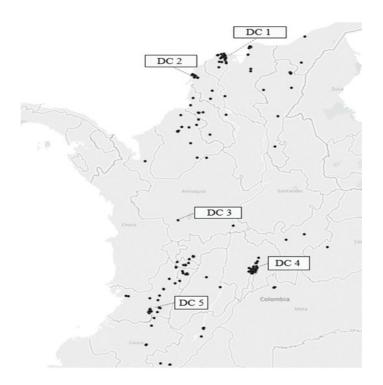
Figure 2. Location of supply nodes (DCs) and demand nodes (Stores).

Results and Sensitivity Analysis

First, the problem was solved without the service level limitations to compare the results with or without equations 7 and 8. In Table 2, it can be seen that the number of distribution centers from one to five was forced, establishing $F_{min} = F_{max}$. Table 2 shows the open facilities in each scenario. Likewise, Table 3 presents the total cost in dollars (objective function). It also presents the following set of metrics: the average distance, the weighted average distance, and the maximum distance to a store (the distance of the furthest store from any distribution center in the supply network).

Table 2. Case study results without the service level restrictions (Equations 7 and 8) – DCs opens.

Target Function: TotalCost Number of variables: 1535; Number of restrictions: 1843					
	DC 1	DC 2	DC 3	DC 4	DC 5
$F_{min} = F_{max} = 1$			Open		
$F_{min} = F_{max} = 2$	Open				Open
$F_{min} = F_{max} = 3$	Open			Open	Open
$F_{min} = F_{max} = 4$	Open	Open		Open	Open
$F_{min} = F_{max} = 5$	Open	Open	Open	Open	Open

Table 3. Case study results without the service level restrictions (Equations 7 and 8).

$F_{min} = F_{max}$	1	2	3	4	5
Total Cost (USD)	$62,339,6	$42,554,9	$46,464,1	$57,126,3	$69,143,7
Average distance (miles)	332.81	130.55	64.13	53.21	51.54
Weighted average distance (miles)	339.94	122.53	65.93	55.15	53.57
Maximum distance to a store (miles)	585	507	349	295	259
% of demand within the range of distance:					
Distance ≤ 10 miles	0.00%	23.75%	31.44%	40.20%	40.20%
10 miles < distance ≤ 25 miles	0.00%	8.50%	18.41%	18.41%	18.41%
25 miles < distance ≤ 50 miles	0.00%	2.85%	3.39%	3.39%	3.39%
50 miles < distance ≤ 100 miles	0.00%	21.60%	22.90%	15.13%	15.13%
100 miles < distance ≤ 150 miles	5.34%	7.14%	8.03%	8.61%	9.91%
150 miles < distance	94.66%	36.16%	15.82%	14.26%	12.95%

* Current company scenario: 5 distribution centers operating in Colombia

Finally, this table presents the percentage of demand supplied within certain distance intervals. In other words, it presents what percentage of demand is within 10 miles, 10 miles to 25 miles, 25 to 50 miles, 50 to 100 miles, 100 to 150 miles, and more than 150 miles. It is important to mention that the solution that includes five distribution centers corresponds to the current scheme managed by the company illustrated in Figure 3. The results show that, in terms of cost, the optimal solution includes two distribution centers ($F_{min} = F_{max} = 2$) with the lowest total cost (USD 42,554,903). However, it can be seen from Table 3 that 36.16% of the demand is distributed over 150 miles, which is a clear strategic disadvantage.
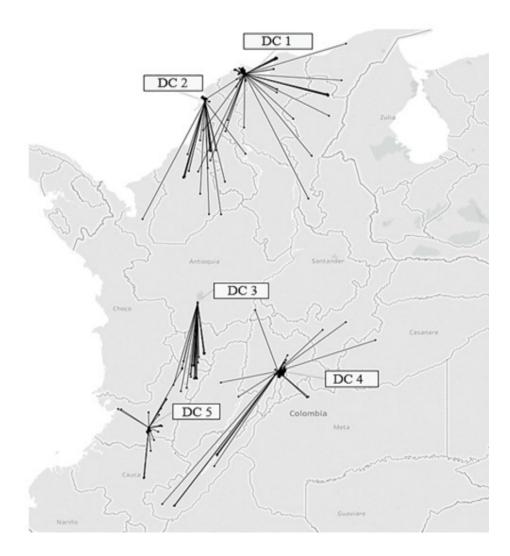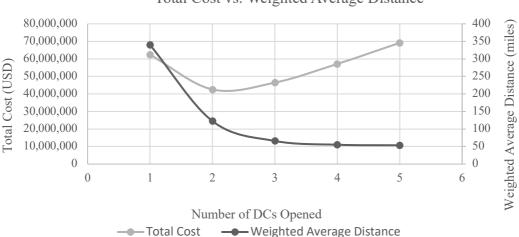
Figure 3. Location of supply nodes (CEDIS) and demand nodes (Shops).

Importantly, the model indicated which facility to open and how much flows in each arc between a supply node and a demand node, and as the number of open facilities increased, the cost could be expected to rise. In turn, as more distribution centers are opened, they will be located closer to the stores, which realistically means that property value will be more expensive because facilities are being in denser areas. Figure 4 shows the number of open distribution centers on the horizontal axis, and the weighted average distance and cost are shown on the vertical axis. As the number of distribution centers increases, the distance to the stores decreases. The cost initially decreases but then begins to increase as facility costs begin to dominate.

Figure 4. Sensitivity analysis - Total Cost vs. Weighted Average Distance.

Figure 5 shows a different metric. The total cost is presented on the left axis, and the right axis shows the percentage of demand supplied within 10 miles, which is between 0% and 40.20% (see Table 3). Again, as the number of distribution centers increases, the level of service increases, because these distribution centers are closer to the stores. However, considering the optimal solution where $F_{min} = F_{max} = 2$, only 23.75% of all demand is within a 10-mile radius, and most demand is over 150 miles away. Therefore, this raises specific questions related to what would happen if the company wants to make sure that its stores are satisfied and when they order something, they get it quickly or where they should locate other facilities to help them increase their level of service.

However, by solving the same case study, but instead of just reporting the distance metrics, the case was solved by adding the service level restrictions (equations 7 and 8). It is important to mention that nothing changed except the inclusion of these restrictions, so the maximum allowable average distance was set at 150 miles. Similarly, the minimum allowable demand within 150 miles was set at less than 85%. Therefore, at least 85% of the demand had to be within 150 miles.
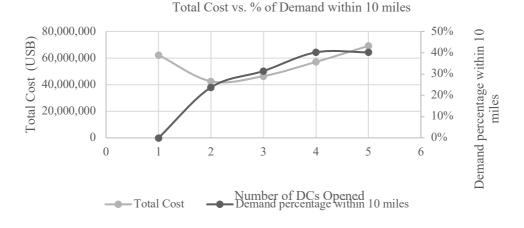

Figure 5. Sensitivity analysis - Total cost vs. demand within 10 miles.

Again, the number of distribution centers was forced to analyze the behavior of the model in various scenarios, $F_{min} = F_{max}$. This is presented in Table 4. Scenarios with only 1, 2, or 3 distribution centers reflected an unfeasible solution in SAS, as they cannot meet service level restrictions.

Table 4. Results of the case study with the service level restrictions in Equations 7 and 8 - DCs opens.

Target Function: TotalCost Number of variables: 1535; Number of restrictions: 1845					
	DC 1	DC 2	DC 3	DC 4	DC 5
$F_{min} = F_{max} = 1$	unfeasible solution.				
$F_{min} = F_{max} = 2$	unfeasible solution.				
$F_{min} = F_{max} = 3$	unfeasible solution.				
$F_{min} = F_{max} = 4$	Open	Open		Open	Open
$F_{min} = F_{max} = 5$	Open	Open	Open	Open	Open

When the problem was rerun in SAS setting $F_{min} = 1$ and $F_{max} = 5$, the optimal solution yielded four distribution centers. The results are presented in Table 5.

Table 5. Results of the case study with the service level restrictions in Equations 7 and 8.

$F_{min} = F_{max}$	4*	5
Total Cost (USD)	$ 57.126.309	$ 69.143.715
Average distance (miles)	53,21	51,54
Weighted average distance (miles)	55,15	53,57
Maximum distance to a store (miles)	295	259
% of demand within the range of distance:		
Distance ≤ 10 miles	40,20%	40,20%
10 miles < distance ≤ 25 miles	18,41%	18,41%
25 miles < distance ≤ 50 miles	3,39%	3,39%
50 miles < distance ≤ 100 miles	15,13%	15,13%
100 miles < distance ≤ 150 miles	8,61%	9,91%
150 miles < distance	14,26%	12,95%

*Optimal solution.

Notably, the distance to the stores has decreased as the company opens more distribution centers. The same happens with the average distance to supply the demand, instead of the 23.75% obtained in Table 3, the new solution presents that 40.20% of the demand of the stores is satisfied at 10 miles. The other results are summarized in Table 5. The scenario with five distribution centers presents similar values; however, it is more expensive. Thus, the optimal configuration is illustrated in Figure 6.

Similarly, when comparing the maximum distance to a store in the two runs, it is required to step value from 507 to 295 miles away, which is a considerable improvement. Also, demand over 150 miles decreased from 36.16% to 14.26%. Therefore, it is remarkable to infer that as the number of distribution centers increases, the service level will tend to improve. On the other hand, costs do not always go in the same direction. Each new facility that the company opens adds more cost, accompanied by a decrease in transportation costs.
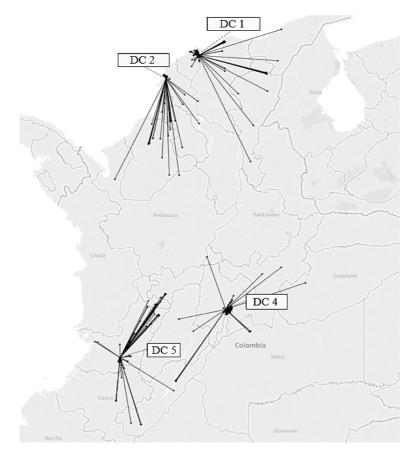

Figure 6. Optimal configuration.

CONCLUSIONS AND DISCUSSION

It was possible to design a model based on mixed-integer linear programming to solve facility location problems, proposing a new methodology that considers the location alternatives that are viable for a company and result from a previous analysis. It is important to highlight that this model can be applicable to many real-life problems considering that as markets emerge in several regions of the world, it might be observed that economic, political, and sociocultural considerations lead to some key factors that managers have to balance in making locational decisions. One factor is the attractiveness of candidate locations as suggested by budget considerations and the size of the prospective market. Others are possible risks concerning political instability, insecurity, economic collapse in some regions, and loss of goodwill. Furthermore, the model was solved with a modern commercial optimizer rather simply. The use of Google Maps in conjunction with the optimization program (SAS) is enormously valuable. In addition, to simplifying the analysis of results, Google Maps is used for computing data for path assignment constraints, using

features such as finding shortest paths and create buffers, or areas of influence, around centers with distances measured on the road network.

It should be underlined that the service constraints associated with distances are one of the most critical aspects of the problem of locating facilities in a logistics network. In an alternative scenario, if the model did not have to consider this aspect at all, it would only open a certain number of distribution centers at a lower cost. In this manner, the results showed that, for higher investment, a company could obtain a better level of service, in addition to certain associated strategic advantages. Just to mention an example related to the case study presented in this paper, the company could backtrack and say that it is not necessary to supply 85% of the demand within 150 miles. Perhaps they could set a lower percentage and, as the problem becomes more complex, it is possible to run different scenarios with these values and analyze what actually increases the cost or if the cost increases dramatically. In a similar way, with regard to the case study, the solution says that it is convenient to open four distribution centers, but perhaps the company does not want to implement that optimal least-cost solution, and actually wants to opt for another strategy. This is a clear advantage of this location model; it is possible to study many hypothetical scenarios.

Besides, the model is useful for analyzing the fact that the cost of supply will decrease if a company has facilities in a distribution network closer to facilities in the retail network. As a company has more and more supply nodes, it will be closer and closer to achieving good results in productivity and service indicators. Retailers generally open more distribution centers near urban areas, to reduce transportation costs and also improve response time. At the same time, as a company has more of these distribution facilities, the cost of transportation from the plants or factories to these facilities increases. There could be two reasons for this: the first one logically relates to the fact that there are more facilities, but the second one is related to the fact that the company has fewer goods going to each of these facilities, in other words, there is a lower aggregation of goods. Then, a suggested solution could be to ship by full truckload, which is the lowest unit cost mode, or to use a transport mode such as LTL (Less Than Truck Load) or multistop truckload This is an aspect that should be investigated in future study cases.

Finally, the model still possesses limitations. The model proposed is quite basic and can be improved in several directions. First, the model could be enhanced to include the production of multiple products in the facility location problem. This would also provide an analysis of the trade-offs between dedicated and flexible manufacturing technologies at the facility level. Second, differentiating characteristics of global manufacturing, such as exchange rates and uncertainties of prices, taxes and quotes, could also be incorporated into integrated facility design models. Third, the model could be extended to include the possible expansion of multiple echelons of facilities, such as manufacturing, processing or assembly and subassembly plants, distribution centers, and warehouses. Fourth, the model could be extended to capture uncertainty in demand and forecast error. Finally, the existence of scale economies in transportation could be included in the proposed model.

APPENDIX A

Table A1. SAS Code.

Line	
1	proc optmodel;
2	set SNodes = {/* Enter the supply nodes*/};
3	set Dnodes = {/* Enter the demand nodes*/};
4	Number distance{SNodes, Dnodes}=[/* Enter the distance matrix*/];

5	Number proximity{SNodes, Dnodes}=[/* Enter the proximity matrix*/];
6	number varcost = /* Enter the variable cost*/;
7	number fixcost{SNodes} = [/*Enter the fixed costs for each facility from the set of supply nodes*/];
8	number capacity{SNodes} = [/*Enter the capacity of each facility from the set of supply nodes*/];
9	number demand{Dnodes} = [/*Enter the quantity demanded in each of the facilities within the set of demand nodes*/];
10	number Fmin = /* Enter minimum number of facilities*/;
11	number Fmax = /* Enter maximum number of facilities*/;
12	number totaldemand = sum{j in Dnodes}demand[j];
13	var open{SNodes} binary;
14	var flow{Snodes, Dnodes} >= 0;
15	minimize TotalCost = sum{i in Snodes}fixcost[i]*open[i] + varcost*sum{i in Snodes, j in Dnodes}flow[i,j]*distance[i,j];
16	con capacitycon{i in Snodes}: sum{j in Dnodes}flow[i,j] <= capacity[i];
17	con demandcon{j in Dnodes}: sum{i in Snodes}flow[i,j] = demand[j];
18	con mincon: sum{i in Snodes}open[i] >= Fmin;
19	con maxcon: sum{i in Snodes}open[i] <= Fmax;
20	con linkingcon{i in Snodes, j in Dnodes}: flow[i,j]-totaldemand*open[i] <= 0;
21	con maxavediscon: sum{i in Snodes, j in Dnodes}distance[i,j]*flow[i,j] <= /* Enter the maximum weighted average distance*/ *totaldemand;
22	con minallowabledemandcon: sum{i in Snodes, j in Dnodes}proximity[i,j]*flow[i,j] >= /* enter the percentage of demand that must be met within a minimum distance value*/ *totaldemand;
23	solve;
24	print TotalCost open flow ;
25	expand;
26	quit;

REFERENCES

Afify, B. et al. 2019. Evolutionary learning algorithm for reliable facility location under disruption. *Expert Systems with Applications* 115: p.223–244.

Biajoli, F.L., Chaves, A.A., & Lorena, L.A.N. 2019. A biased random-key genetic algorithm for the two-stage capacitated facility location problem. *Expert Systems with Applications* 115: p.418–426.

Boujelben, M.K., Gicquel, C., & Minoux, M. 2016. A MILP model and heuristic approach for facility location under multiple operational constraints. *Computers & Industrial Engineering* 98: p.446–461.

Branson, S. et al. 2018. From Google Maps to a fine-grained catalog of street trees. *ISPRS Journal of Photogrammetry and Remote Sensing* 135: p.13–30.

Cadarso, L., Codina, E., Escudero, L.F., & Marín, A. 2017. Rapid transit network design: Considering recovery robustness and risk aversion measures. *Transportation Research Procedia* 22: p.255–264.

Catal, C., Fenerci, A., Ozdemir, B., & Gulmez, O. 2015. Improvement of Demand Forecasting Models with Special Days. *Procedia Computer Science* 59: p.262–267.

Cedolin, M., Göker, N., Dogu, E., & Esra Albayrak, Y. 2018. Facility Location Selection Employing Fuzzy DEA and Fuzzy Goal Programming Techniques. In J. Kacprzyk, E. Szmidt, S. Zadrożny, K. T. Atanassov, & M. Krawczak (eds) *Advances in Fuzzy Logic and Technology 2017*, 466–476. Cham: Springer International Publishing

Chalupa, D., & Nielsen, P. 2019. A simple and robust Monte Carlo hybrid local search algorithm for the facility location problem. *Engineering Optimization* 51(5): p.832–845.

Chopra, S., & Meindl, P. 2016. *Supply Chain Management: Strategy, Planning, and Operation. Chapter 1: Understanding the Supply Chain. Section 1.4 Decision phases in a supply chain* 6th ed. Pearson Education (ed). USA.

Coniglio, S., Fliege, J., & Walton, R. 2017. Facility Location with Item Storage and Delivery. In A. Sforza & C. Sterle (eds) *Optimization and Decision Science: Methodologies and Applications*, 287–294. Cham: Springer International Publishing

Correia, I., & Melo, T. 2016. Multi-period capacitated facility location under delayed demand satisfaction. *European Journal of Operational Research* 255(3): p.729–746.

Correia, I., & Melo, T. 2017. A multi-period facility location problem with modular capacity adjustments and flexible demand fulfillment. *Computers & Industrial Engineering* 110: p.307–321.

Council of Supply Chain Management Professionals (CSCMP). 2013. Supply chain management: Terms and Glossary. : p.222. Available at: https://cscmp.org/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms/CSCMP/Educate/.

Ding, L., & Zhang, N. 2016. A Travel Mode Choice Model Using Individual Grouping Based on Cluster Analysis. *Procedia Engineering* 137: p.786–795.

Drezner, T., Drezner, Z., & Schöbel, A. 2018. The Weber obnoxious facility location model: A Big Arc Small Arc approach. *Computers & Operations Research* 98: p.240–250.

Du, B., Zhou, H., & Leus, R. 2020. A two-stage robust model for a reliable p-center facility location problem. *Applied Mathematical Modelling* 77: p.99–114.

Emirhüseyinoğlu, G., & Ekici, A. 2019. Dynamic facility location with supplier selection under quantity discount. *Computers & Industrial Engineering* 134: p.64–74.

Farahani, R.Z., Fallah, S., Ruiz, R., Hosseini, S., & Asgari, N. 2019. OR models in urban service facility location: A critical review of applications and future developments. *European Journal of Operational Research* 276(1): p.1–27.

Golpîra, H. 2020. Optimal integration of the facility location problem into the multi-project multi-supplier multi-resource Construction Supply Chain network design under the vendor managed inventory strategy. *Expert Systems with Applications* 139: p.112841.

Google LLC. 2017a. Distance Matrix API: Developer Guide. *Google Maps Platform*. Available at: https://developers.google.com/maps/documentation/distance-matrix/intro [Accessed January 10, 2017].

Google LLC. 2017b. Google Maps Platform. *Google Cloud*. Available at: https://cloud.google.com/maps-platform/maps/ [Accessed January 10, 2017].

Guo, P., Cheng, W., & Wang, Y. 2017. Hybrid evolutionary algorithm with extreme machine learning fitness function evaluation for two-stage capacitated facility location problems. *Expert Systems with Applications* 71: p.57–68.

Hajipour, V., Fattahi, P., Tavana, M., & Caprio, D. Di. 2016. Multi-objective multi-layer congested facility location-allocation problem optimization with Pareto-based meta-heuristics. *Applied Mathematical Modelling* 40(7): p.4948–4969.

Jakubovskis, A. 2017. Strategic facility location, capacity acquisition, and technology choice decisions under demand uncertainty: Robust vs. non-robust optimization approaches. *European Journal of Operational Research* 260(3): p.1095–1104.

Karatas, M. 2017. A multi-objective facility location problem in the presence of variable gradual coverage performance and cooperative cover. *European Journal of Operational Research* 262(3): p.1040–1051.

Karatas, M., & Yakıcı, E. 2018. An iterative solution approach to a multi-objective facility location problem. *Applied Soft Computing* 62: p.272–287.

Khosravi, S., & Jokar, M.R.A. 2017. Facility and hub location model based on gravity rule. *Computers & Industrial Engineering* 109: p.28–38.

Kınay, Ö.B., Saldanha-da-Gama, F., & Kara, B.Y. 2019. On multi-criteria chance-constrained capacitated single-source discrete facility location problems. *Omega* 83: p.107–122.

Kumarage, S. 2018. *Use of Crowdsourced Travel Time Data in Traffic Engineering Applications*. University of Moratuwa.

Land, A.H., & Doig, A.G. 1960. An Automatic Method of Solving Discrete Programming Problems. *Econometrica* 28(3): p.497–520.

Lin, Y. et al. 2018. An improved artificial bee colony for facility location allocation problem of end-of-life vehicles recovery network. *Journal of Cleaner Production* 205: p.134–144.

Manthey, B., & Tijink, M.B. 2018. Perturbation resilience for the facility location problem. *Operations Research Letters* 46(2): p.215–218.

Mei, L., Li, M., Ye, D., & Zhang, G. 2019. Facility location games with distinct desires. *Discrete Applied Mathematics* 264: p.148–160.

Nasiri, M.M., Mahmoodian, V., Rahbari, A., & Farahmand, S. 2018. A modified genetic algorithm for the capacitated competitive facility location problem with the partial demand satisfaction. *Computers & Industrial Engineering* 124: p.435–448.

Orjuela-Castro, J.A., Sanabria-Coronado, L.A., & Peralta-Lozano, A.M. 2017. Coupling facility location models in the supply chain of perishable fruits. *Research in Transportation Business & Management* 24: p.73–80.

Pusponegoro, N.H., Rachmawati, R.N., Notodiputro, K.A., & Sartono, B. 2017. Linear Mixed Model for Analyzing Longitudinal Data: A Simulation Study of Children Growth Differences. *Procedia Computer Science* 116: p.284–291.

Qi, M., Xia, M., Zhang, Y., & Miao, L. 2017. Competitive facility location problem with foresight considering service distance limitations. *Computers & Industrial Engineering* 112: p.483–491.

Rachmawati, R.N., & Pusponegoro, N.H. 2017. Hierarchical Linear Mixed Model for Poverty Analysis in Indonesia. *Procedia Computer Science* 116: p.182–189.

Rohaninejad, M., Navidi, H., Nouri, B.V., & Kamranrad, R. 2017. A new approach to cooperative competition in facility location problems: Mathematical formulations and an approximation algorithm. *Computers & Operations Research* 83: p.45–53.

SAS Institute Inc. 2014. *SAS/OR® 13.2 User's Guide: Mathematical Programming - The OPTMODEL Procedure*. North Carolina, USA.

SAS Institute Inc. 2018. SAS/OR(R) 14.1 User's Guide: Mathematical Programming - The Mixed Integer Linear Programming Solver. Available at: http://support.sas.com/documentation/cdl/en/ormpug/68156/HTML/default/viewer.htm#ormpug_milpsolver_overview.htm [Accessed January 10, 2018].

Sauvey, C., Melo, T., & Correia, I. 2019. Heuristics for a multi-period facility location problem with delayed demand satisfaction. *Computers & Industrial Engineering*: p.106171.

Semeida, A.M. 2014. Derivation of travel demand forecasting models for low population areas: the case of Port Said Governorate, North East Egypt. *Journal of Traffic and Transportation Engineering (English Edition)* 1(3): p.196–208.

Tadros, S.A., Galal, N.M., Ghazy, M., & ElSayed, A.E. 2018. A multi-objective two echelon capacitated facility location problem. In *2018 7th International Conference on Industrial Technology and Management (ICITM)*, 260–264.

Teye, C., Bell, M.G.H., & Bliemer, M.C.J. 2017. Urban intermodal terminals: The entropy maximising facility location problem. *Transportation Research Part B: Methodological* 100: p.64–81.

Wang, Wei, Wang, Wanmei, Mosley, T.H., & Griswold, M.E. 2017. A SAS macro for the joint modeling of longitudinal outcomes and multiple competing risk dropouts. *Computer Methods and Programs in Biomedicine* 138: p.23–30.

Yang, S.Y., & Hsu, C.L. 2016. A location-based services and Google maps-based information master system for tour guiding. *Computers and Electrical Engineering* 54: p.87–105.

Yang, Z., Chen, H., Chu, F., & Wang, N. 2019. An effective hybrid approach to the two-stage capacitated facility location problem. *European Journal of Operational Research* 275(2): p.467–480.

Yu, G., Haskell, W.B., & Liu, Y. 2017. Resilient facility location against the risk of disruptions. *Transportation Research Part B: Methodological* 104: p.82–105.