

## CAPACITATED FACILITY LOCATION PROBLEM WITH GENERAL OPERATING AND BUILDING COSTS

**José A. Montoya**

Departamento de Ingeniería de Producción, Universidad EAFIT  
Carrera 49 No. 7 sur 50, Medellín, Colombia  
[jmonto36@eafit.edu.co](mailto:jmonto36@eafit.edu.co)

**Mario C. Vélez-Gallego**

Departamento de Ingeniería de Producción, Universidad EAFIT  
Carrera 49 No. 7 sur 50, Medellín, Colombia  
[marvelez@eafit.edu.co](mailto:marvelez@eafit.edu.co)

**Juan G. Villegas**

Departamento de Ingeniería Industrial, Universidad de Antioquia  
Calle 70 No. 52 – 21, 050010 – Medellín, Colombia  
[jvillega@udea.edu.co](mailto:jvillega@udea.edu.co)

### ABSTRACT

This research introduces the capacitated facility location problem with general operating and building costs (CFLPGOBC). The CFLPGOBC extends previous problems found in the literature by including general operating and building cost functions that allows modeling different behaviors (i.e. economies of scale, congestion, etc.). The CFLPGOBC was formulated as a mixed integer linear program (MILP) and solved using a commercial solver. The performance of the proposed formulation was evaluated with a set of randomly generated test instances. After one hour of computational time the solver converged to an optimality gap of 1% or less in 55% of the instances tested. The overall optimality gap was 3.57%, on average.

**KEYWORDS.** Facility location. Mixed integer linear programming. Supply chain design

**L & T - Logistics and Transport**

**PM - Mathematical Programming**

**IND - OR in Industry**

## 1. Introduction

The Capacitated Facility Location Problem (CFLP) is one of the most widely studied problems in supply chain design. The problem consists of i) determining the quantity and location of a set of capacitated facilities which will produce/distribute a given product or service and ii) allocating the amount of product to ship from each facility to a set of customers with the goal of optimizing a given objective function, which is frequently a total relevant cost function. The classical CFLP assumes a fixed cost to open a facility and a unitary cost of shipping products from a given facility to a customer location. However, in many real-world applications the cost of opening a facility is a non-linear function of the size of the facility. Among others, one reason for this non-linear behavior is the fact that machine or plant capacities can only be increased by discrete amounts. Additionally, some of the models available in the literature do not consider the cost of operating a facility into the total relevant cost function. The absence of this cost implies the assumption that the cost of operating a facility is independent of the site where the facility is to be opened and its size, which is not a reasonable assumption in many cases. Furthermore and similar to the opening cost, the operating cost function may be non-linear with the size of the facility. This non-linear behavior may be caused, among other reasons, by economies or diseconomies of scale and congestion phenomena. To bridge this gap, this research introduces the so-called capacitated facility location problem with general operating and building costs (CFLPGOBC) which is inspired by the structure of the Colombian cement industry. A mixed integer linear program (MILP) is introduced to solve the CFLPGOBC. Extensive computational experiments were conducted in order to assess the performance of the proposed formulation when implemented in a commercial solver. This paper is organized as follows. Section 2 presents the literature review of the CFLP and other related problems. In section 3 the problem is formulated as a MILP. A summary of the computational experiments conducted to assess the performance of the proposed formulation is presented in section 4. Conclusions and future research are presented in section 5.

## 2. Literature Review

The Facility Location Problem (FLP) is a classical operations research problem that has been addressed by numerous researchers over the past few years (see (M. S. Daskin 2008), (C. S. ReVelle et al. 2008) and references therein). Two typical variants of the problem are clearly identified in the literature: The uncapacitated facility location problem (UFLP) and the capacitated facility location problem (CFLP). Both the UFLP and the CFLP consist of deciding where to open a set of facilities and how customers should be assigned to these facilities so that a total cost function is minimized. The difference between both is that in the UFLP the facilities do not have a constraint imposed on the maximum capacity while in the CFLP there is a capacity constraint on the maximum demand that can be assigned to each open facility. Both the UFLP and the CFLP are shown to be NP-hard (Cornuejols et al. 1990) and have been extensively studied.

Several variants of the CFLP have been studied. In (Melo et al. 2009) the authors presented an exhaustive review of the problem, its extensions, solution methods and real-world applications. The most important extensions to this problem involve single or multiple time periods, deterministic or stochastic demand, single or multiple products, and single or multiple layers, among others. In (Current et al. 2002) the authors review several applications of facility location models that include banks, airports, bus stops, fast foods, plants and warehouses. In general the CFLP has been solved by both exact and approximate methods including branch-and-bound (Akinc & Khumawala 1977), Lagrangian relaxation (Christofides & Beasley 1983), Benders decomposition (Geoffrion & Graves 1974), genetic algorithms (Kratika et al. 2001), tabu search (Sun 2012), artificial neural networks (Vaithyanathan et al. 1996), simulated annealing (Arostegui Jr. et al. 2006) and greedy randomized adaptive search procedures (Resende & Werneck 2006) among others.

Some researchers have worked on the CFLP with different cost structures (s-shape,

staircase, convex, concave and general). These problems have been solved with different methods (Lagrangian relaxation, column generation, Benders decomposition and metaheuristics). Table 1 presents a comparison of the solution approaches and cost functions considered in the most important references addressing the CFLP.

**Table 1. Classification of some important references based on the cost function**

References	Function	Method †
(Ozsen et al. 2009), (Schütz et al. 2008), (Broek et al. 2006)	S shape	LR
(Holmberg & Ling 1997)	Staircase	LR
(Wu et al. 2006)	General	LR
(Desrochers et al. 1995)	Convex	CG
(Cohen & Moon 1991)	Concave	BD
(Dupont 2008), (Dasci & Verter 2001)	Concave	BB
(Harkness & C. ReVelle 2003)	Convex	BB
(Romeijn et al. 2010), (Lin et al. 2006), (Hajiaghayi et al. 2003)	Concave	MH

† LR: Lagrangian Relaxation, CG: Column Generation, BD: Benders Decomposition, BB: Branch & Bound, MH: Metaheuristic

To the best of our knowledge the only two papers that address problems similar to the CFLP<sub>GOBC</sub> are (Wu et al. 2006) and (Dupont 2008). (Wu et al. 2006) formulated a CFLP with a general setup cost. In this work the authors considered a fixed cost to open a facility, a shipping unitary cost, and a general setup cost which depends on the site, the type of facility and the capacity of the facility. They assumed that in a given site several facilities can be installed, each for a different product. The most important differences between the problem addressed in (Wu et al. 2006) and the CFLP<sub>GOBC</sub> is that in the cement industry as in many other industries, in one single facility several different products can be produced, therefore the building cost of a facility depends only on the capacity of the facility and the site (i.e. land cost), and usually it is independent from the exact mix of products that the facility can produce. On the other hand, the operating cost depends on the site (i.e. cost of the resources) and on the amount of each product allocated to the facility to produce. For that reason, in our case it is more convenient to separate the operating and building costs in two independent terms within the total cost function. In (Dupont 2008) the author introduces a similar problem based on real-world applications where the building, operating, and shipping costs are concave functions that depend on the quantity of the product and the site. In the MILP he proposed these costs are expressed in a global cost function for each site. The CFLP<sub>GOBC</sub> extends the problem proposed in (Dupont 2008) by including more general cost functions, capacitated facilities and multiple products.

### 3. Mathematical Model

To model the CFLP<sub>GOBC</sub> as a mixed integer linear program (MILP) we use the following notation: Let  $N=\{1, \dots, n\}$  be the set of candidate locations at which a facility can be built, the quantity  $u_i$  represents an upper bound on the capacity (i.e. size) of the facility that can be built at location  $i \in N$ . The set  $Q=\{1, \dots, U\}$  states the feasible production quantities and  $L=\{1, \dots, U\}$  the set of feasible sizes for the facilities, where  $U = \max_{i \in N}\{u_i\}$ .

Let  $M=\{1, \dots, m\}$  be the set of customers and  $P=\{1, \dots, p\}$  the set of products. Each customer  $j \in M$  requires  $d_{jk}$  units of product of type  $k \in P$  (i.e. demand), and  $c_{ijk}$  is the cost of transporting one unit of product  $k \in P$  from location  $i \in N$  to customer  $j \in M$ . The cost of building a facility of capacity  $l \in L$  at candidate location  $i \in N$  is given by  $f_{il}$ . Whereas  $g_{ikq}$  represents the cost of producing the quantity  $q \in Q$  of product  $k \in P$  in a facility located at  $i \in N$ . To model the different decisions addressed in the CFLP<sub>GOBC</sub> we use three sets of decision variables:

$x_{ijk}$ : The fraction of demand of product  $k$  supplied to customer  $j$  from facility  $i$ .

$y_{il}$ : Binary variable that indicates the size of facility  $i$ .  $y_{il}$  takes the value of 1 if  $l$  is the size of facility  $i$  and 0 otherwise.

$z_{ikq}$ : Binary variable that indicates the quantity of item  $k$  produced at facility  $i$ .  $z_{ikq}$  takes the value of 1 if  $q$  is the quantity produced of item  $k$  at facility  $i$  and 0 otherwise.

Using the above notation the CFLPGOBC can be formulated as follows:

$$\text{Min} \sum_{i \in N} \sum_{l \in L} f_{il} \cdot y_{il} + \sum_{i \in N} \sum_{k \in P} \sum_{q \in Q} g_{ikq} \cdot z_{ikq} + \sum_{i \in N} \sum_{j \in M} \sum_{k \in P} c_{ijk} \cdot x_{ijk} \quad (1)$$

Subject to:

$$\sum_{i \in N} x_{ijk} = 1 \quad \forall j \in M, \forall k \in P \quad (2)$$

$$\sum_{j \in M} d_{jk} \cdot x_{ijk} \leq \sum_{q \in Q} q \cdot z_{ikq} \quad \forall i \in N, \forall k \in P \quad (3)$$

$$\sum_{k \in P} \sum_{q \in Q} q \cdot z_{ikq} \leq \sum_{l \in L} l \cdot y_{il} \quad \forall i \in N \quad (4)$$

$$\sum_{l \in L} l \cdot y_{il} \leq u_i \quad \forall i \in N \quad (5)$$

$$\sum_{l \in L} y_{il} \leq 1 \quad \forall i \in N \quad (6)$$

$$\sum_{q \in Q} z_{ikq} \leq 1 \quad \forall i \in N, \forall k \in P \quad (7)$$

$$x_{ijk} \geq 0 \quad \forall i \in N, \forall j \in M, \forall k \in P \quad (8)$$

$$y_{il} \in \{0,1\} \quad \forall i \in N, \forall l \in L \quad (9)$$

$$z_{ikq} \in \{0,1\} \quad \forall i \in N, \forall k \in P, \forall q \in Q \quad (10)$$

The objective function (1) is to minimize the total cost function: transportation, operating and building costs. Constraint set (2) ensures that the demand of each customer is satisfied; constraint set (3) states that the demand supplied by facility  $i$  of product  $k$  cannot exceed the quantity of product  $k$  produced by the facility. Constraint set (4) ensures that the quantity produced in each facility does not exceed the size of the facility. Constraint set (5) ensures that the size of each facility does not exceed the maximum capacity of the site; constraint set (6) establishes that each facility can only take one value of size; constraint set (7) establishes that for each facility, only one level of production is selected for each product. Constraint sets (8), (9) and (10) define the domain of the decision variables.

Using the above formulation, it can be seen that the CFLPGOBC is NP-hard since it generalizes the CFLP (when  $k=1$ ,  $l=1$ , and  $g_{ikq}=0 \forall i \in N$ ), which is an NP-hard problem (Cornuejols et al. 1990).

#### 4. Computational Experiments

A set of computational experiments were executed to evaluate the performance of the mathematical model. The CFLPGOBC was implemented in Matlab 7.12 and Gurobi 4.5.1 was used to solve the MILP. The connection between Gurobi and Matlab was made using the Gurobi Mex interface (Yin 2012). The optimality tolerance for Gurobi was set at  $10^{-6}$  and the solver was allowed to run for a maximum of 3600 seconds. The experiments were run on a 3.07 GHz Intel

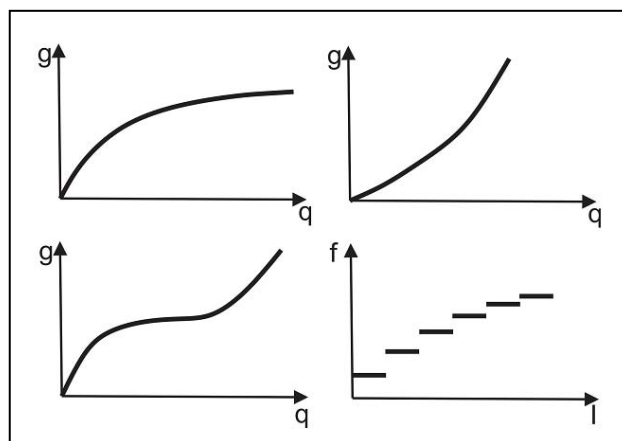
Core i7 with 16 GB of memory running Window 7 at 64 bits.

The test instances were randomly generated based on information from the Colombian cement industry. The sizes and localizations of the candidate locations were taken from the actual plants and possible calyx mines. The sizes and locations of the customers were taken from the configuration of the cement supply chain, where normally the customers are distributions centers and depots, which are located in main cities. The values of maximal capacities are inspired by the capacity of the actual plants (the values are measured in  $10^4$  tons). The set of products was based on the cement market. The demand and the cost magnitudes were generated having into account the data of a cement company. Table 2 summarizes the parameters used in the generation of the test instances. Three instances were generated for each possible combination of the parameters, for a total of 288 test instances.

**Table 2. Parameters for the generation of the test instances**

Parameters	Notation	Values
Building cost	-	Staircase
Operating cost	-	Concave, Convex, S-shape
Facilities	$n$	10, 20
Customers	$m$	100, 150
Products	$p$	4, 8
Maximum capacity	$U$	100, 200
Demand to capacity ratio	$g$	0.25, 0.75

Since one of the main properties of the CFLPGOBC is the possibility to model any function to represent the operating and building costs, for the generation of the instances three different options were considered for the operating cost: a concave function, a convex function, and an s-shaped function. These options represent respectively: economies of scale, diseconomies of scale and economies of scale at the beginning and diseconomies at the end due to congestion. The building cost function used for all the instances was a staircase function, with the property that when the size of the facilities increases the difference between two consecutive fixed cost decreases. This structure represents economies of scale (Figure 1).



**Figure 1. Functions used in the set of instances (concave, convex, s-shape and staircase respectively)**

For each instance, the upper (UB) and lower bound (LB) obtained by the solver after one hour of computational time were used to compute the optimality gap as in Equation 11. Table 3 summarizes the average gap, number of instances solved to optimality and average run time for each shape of the operating cost function.

$$\%Gap = \frac{UB - LB}{UB} \times 100\% \tag{11}$$

**Table 3. Average gap by operating cost function**

Operating cost function	%Gap		Number of instances solved to optimality	Average run time (s)
	Average	Maximum		
Concave	1.17	10.44	30/96	2728.43
Convex	0.53	4.71	19/96	3044.82
S-shape	9.00	69.51	14/96	3121.54
Overall	3.57	69.51	63/288	2964.93

After one hour of computational time, the average optimality gap was 3.57%, and 55% of the instances were solved with a gap of less than 1%. Table 4 shows gap frequencies for all the instances categorized by the operating cost function. This table shows that for the concave operating cost function 98% of the instances were solved with a gap of less than 10%. In the case of the convex operating cost function 100% of the instances were solved with a gap of less than 5%, and for the s-shape function 67% of the instances achieved a gap of less than 10%. In the 0-1 interval there are significant differences between the number of instances solved to optimality with concave and convex functions as compared to those instances with an s-shape function: only 17% of the instances with as s-shape function were solved with a gap in the 0-1 interval.

**Table 4. Analysis of frequencies for the %Gap**

%Gap Interval	Concave		Convex		S-shape	
	Freq (%)	Cum (%)	Freq (%)	Cum (%)	Freq (%)	Cum (%)
0-1	63	63	86	86	17	17
1-5	34	97	14	100	21	38
5-10	1	98	0	100	29	67
10-20	2	100	0	100	26	93
20-30	0	100	0	100	3	96
30-40	0	100	0	100	2	98
40-50	0	100	0	100	0	98
50-60	0	100	0	100	1	99
60-70	0	100	0	100	1	100

In order to analyze the impact of the characteristics of the problem (i.e. number of candidate facilities, number of customers, maximum capacity and demand to capacity ratio) we analyzed the behavior of the optimality gap for each factor and operating cost function. Table 5 presents the average results obtained for each characteristic for the different operating cost functions, while charts from Figure 2 to Figure 6 summarize the results of these analyses. Note that in the generation of the instances the building cost function was fixed to have a staircase structure.

**Table 5. Average results obtained for each characteristic for the different operating cost functions**

Problem characteristic	Symbol	Levels	Average %Gap			
			Concave	Convex	S-shape	Average
Number of facilities	$m$	10	0.58	0.32	5.70	2.20
		20	1.75	0.73	12.30	4.93
Number of customers	$n$	100	0.87	0.43	6.59	2.63
		150	1.45	0.62	11.41	4.49
Number of products	$p$	4	0.37	0.10	4.80	1.76
		8	1.96	0.94	13.20	5.37
Maximum capacity	$U$	100	1.17	0.79	6.75	2.90
		200	1.16	0.26	11.25	4.22
Demand to capacity ratio	$g$	0.25	0.68	0.35	7.40	2.81
		0.75	1.64	0.70	10.60	4.31

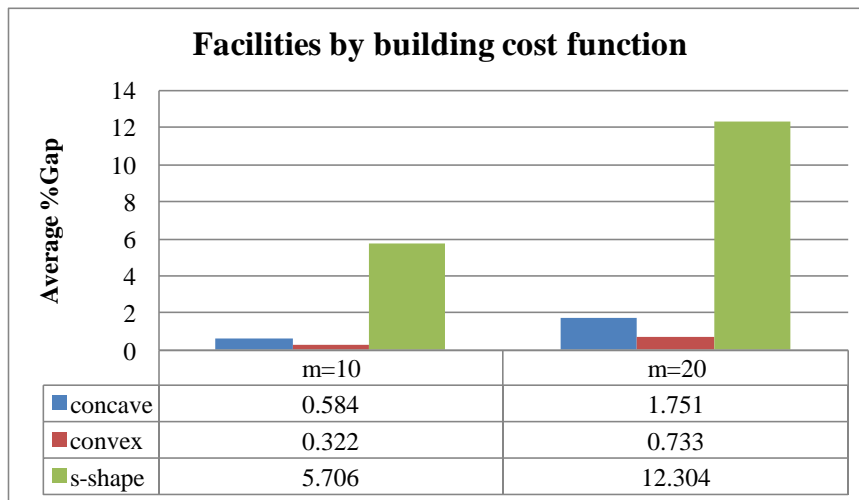


Figure 2. Average optimality gap (%Gap) by number of candidate facilities

As it can be seen in Figure 2, there is a significant increase of the gaps when the number of candidate facilities grows from 10 to 20. Figure 3 presents the behavior when the number of customers changes, once again when the number of customers increases, the optimality gaps increase as well and for the s-shape operating cost function, the impact seems higher. Figure 4 shows that the number of products has an important impact in the behavior of the optimality gaps. If the problem instances have more products; the gaps increase. On the other hand, in Figure 5 it does not seem that when the maximum capacity increases the gaps increase. Finally, Figure 6 shows that the optimality gaps are less sensible to the change in the demand to capacity ratio.

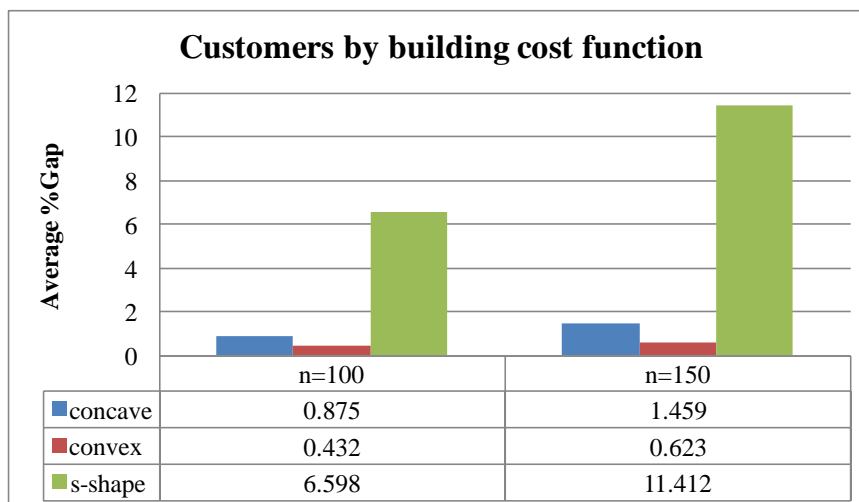


Figure 3. Average optimality gap (%Gap) by number of customers

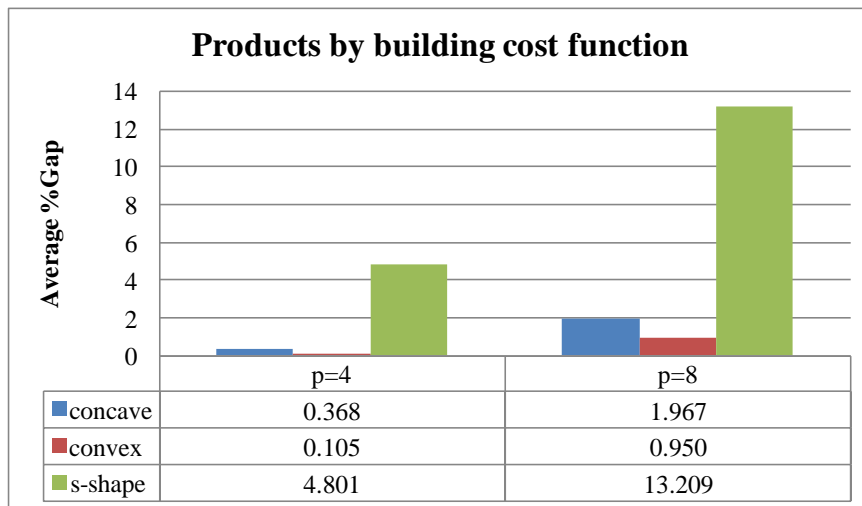


Figure 4. Average optimality gap (%Gap) by number of products

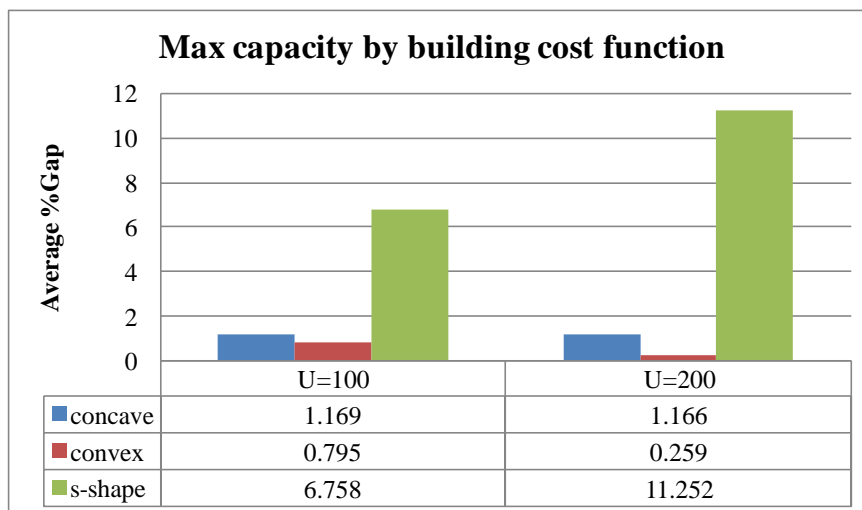


Figure 5. Average optimality gap (%Gap) by maximum capacity

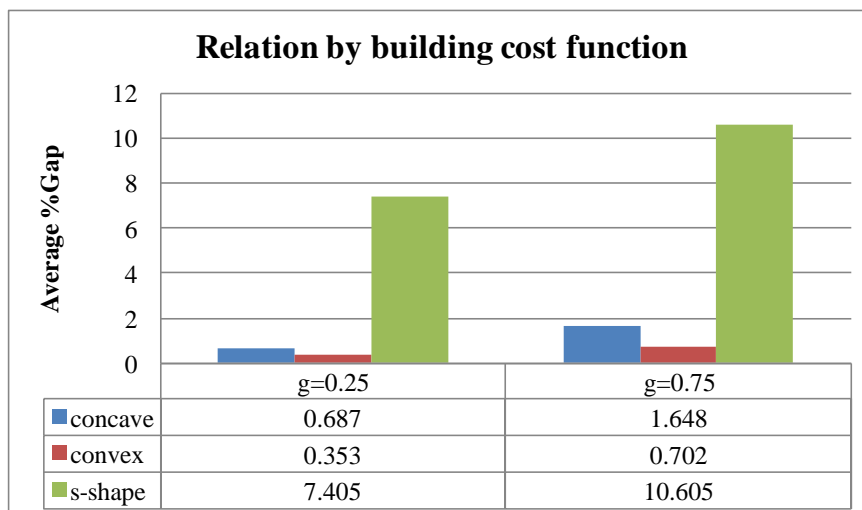


Figure 6. Average optimality gap (%Gap) by demand to capacity ratio

As a summary, the results of this section show that the instances with operating cost



function that follow an s-shape structure have higher optimality gaps. On the other hand instances with convex and concave functions can be solved with consistently smaller optimality gaps.

## 5. Conclusions

This research introduces the capacitated facility location problem with general building and operating costs (CFLPGOBC), a new extension to the capacitated facility location problem with multiple products. The CFLPGOBC has been formulated as a mixed integer linear programming problem that enables the modeling of several building and operating cost functions to represent different behaviors that can be found in practical applications. Even though it was inspired by the structure of the Colombian cement industry which presents economies of scale with a convex operating cost function, the CFLPGOBC allows to model diseconomies of scale (i.e. concave operating cost functions) and other even more complex structures like s-shape and staircase functions.

In addition to the classical decisions considered in facility location problems: the opening of the facilities and the allocation of the customers to the open facilities; the CFLPGOBC also includes two other important decisions that appear in multiproduct environments with (dis)economies of scale: the size of the open facilities (i.e. their total capacities) and the allocation of the total capacity of each facility to each product (i.e. the quantity of each item produced by each facility). In this way the CFLPGOBC allows practitioners to answer two important questions that arise in the design of multi-product supply chain i) is it better to have many small facilities or few big facilities? and ii) is it better to have specialized facilities by product or to have non-specialized multiproduct facilities?

A computational experiment over a test bed of 288 randomly generated instances resembling the structure of the Colombian cement industry was performed in order to test the proposed MILP formulation on a commercial solver. On average, the optimality gap was 3.57%, and 55% of the test instances achieved a gap of less than 1% after one hour of computational time. The computational results revealed that the type of operating cost function has an important impact on the performance of the model. While concave and convex cost functions achieved an average gap of 1.17% and 0.53%, respectively, the test instances with an s-shaped operating cost presented an average gap of 9%. In some of these cases the gaps were significantly high. Future extensions to this research include the evaluation of approximation and decomposition algorithms that allow for improving the convergence of these particularly difficult instances.

## 6. References

- Akinc, U. and Khumawala, B.M.** (1977), An efficient branch and bound algorithm for the capacitated warehouse location problem, *Management Science*, 23(6), 585-594.
- Arostegui Jr., M.A., Kadipasaoglu, S.N. and Khumawala, B.M.** (2006). An empirical comparison of tabu search, simulated annealing, and genetic algorithms for facilities location problems, *International Journal of Production Economics*, 103(2), 742-754.
- Broek, J. v. d., Schütz, P., Stougie, L. and Tomasgard, A.** (2006), Location of slaughterhouses under economies of scale, *European Journal of Operational Research*, 175(2), 740-750.
- Christofides, N. and Beasley, J.E.** (1983), Extensions to a Lagrangean relaxation approach for the capacitated warehouse location problem, *European Journal of Operational Research*, 12(1), 19-28.
- Cohen, M.A. and Moon, S.** (1991), An integrated plant loading model with economies of scale and scope, *European Journal of Operational Research*, 50(3), 266-279.
- Cornuejols, G., Nemhauser, G.L. and Wolsey, L.A.**, The uncapacitated facility location problem, in P. B. Mirchandani and R. L. Francis. (Eds.), *Discrete location theory*, John Wiley & Sons, 119-172, 1990.
- Current, J., Daskin, M. and Schilling, D.**, Discrete network location models, in Z. Drezner and H. W. Hamacher. (Eds), *Facility Location: Applications and Theory*, Springer-Verlag, Berlin, 81-108, 2002

- Dasci, A. and Verter, V.** (2001), The plant location and technology acquisition problem, *IIE Transactions*, 33(11), 963-974.
- Daskin, M.S.** (2008), What you should know about location modeling, *Naval Research Logistics*, 55(4), 283-294.
- Desrochers, M., Marcotte, P. and Stan, M.** (1995), The congested facility location problem, *Location Science*, 3(1), 9-23.
- Dupont, L.** (2008), Branch and bound algorithm for a facility location problem with concave site dependent costs, *International Journal of Production Economics*, 112(1), 245-254.
- Geoffrion, A.M. and Graves, G.W.** (1974), Multicommodity distribution system design by Benders decomposition, *Management Science*, 20(5), 822-844.
- Hajiaghayi, M.T., Mahdian, M. and Mirrokni, V.S.** (2003), The facility location problem with general cost functions, *Networks*, 42(1), 42-47.
- Harkness, J. and ReVelle, C.** (2003), Facility location with increasing production costs, *European Journal of Operational Research*, 145(1), 1-13.
- Holmberg, K. and Ling, J.** (1997), A Lagrangean heuristic for the facility location problem with staircase costs, *European Journal of Operational Research*, 97(1), 63-74.
- Kratka, J., Tasic, D., Filipovic, V. and Liubic, I.** (2001), Solving the simple plant location problem by genetic algorithm, *RAIRO Operations Research*, 35, 127-142.
- Lin, J.-R., Nozick, L. and Turnquist, M.** (2006), Strategic design of distribution systems with economies of scale in transportation, *Annals of Operations Research*, 144(1), 161-180.
- Melo, M.T., Nickel, S. and Saldanha-da-Gama, F.** (2009), Facility location and supply chain management – A review, *European Journal of Operational Research*, 196(2), 401-412.
- Ozsen, L., Daskin, M.S. and Coullard, C.R.** (2009), Facility location modeling and inventory management with multisourcing, *Transportation Science*, 43(4), 455-472.
- ReVelle, C.S., Eiselt, H.A. and Daskin, M.S.** (2008), A bibliography for some fundamental problem categories in discrete location science, *European Journal of Operational Research*, 184(3), 817-848.
- Resende, M.G.C. and Werneck, R.F.** (2006), A hybrid multistart heuristic for the uncapacitated facility location problem, *European Journal of Operational Research*, 174(1), 54-68.
- Romeijn, H. E., Sharkey, T.C., Z-J. M. Shen and Zhang, J.** (2010), Integrating facility location and production planning decisions, *Networks*, 55(2), 78-89.
- Schütz, P., Stougie, L. and Tomasgard, A.** (2008), Stochastic facility location with general long-run costs and convex short-run costs, *Computers & Operations Research*, 35(9), 2988-3000.
- Sun, M.** (2012), A tabu search heuristic procedure for the capacitated facility location problem, *Journal of Heuristics*, 18(1), 91-118.
- Vaithyanathan, S., Burke, L.I. and Magent, M.A.** (1996), Massively parallel analog tabu search using neural networks applied to simple plant location problems, *European Journal of Operational Research*, 93(2), 317-330.
- Wu, L.-Y., Zhang, X.-S. and Zhang, J.-L.** (2006), Capacitated facility location problem with general setup cost, *Computers & Operations Research*, 33(5), 1226-1241.
- Yin, W.**, Gurobi Mex: A MATLAB interface for Gurobi, URL: [http://convexoptimization.com/wikimization/index.php/gurobi\\_mex](http://convexoptimization.com/wikimization/index.php/gurobi_mex), 2009-2011.