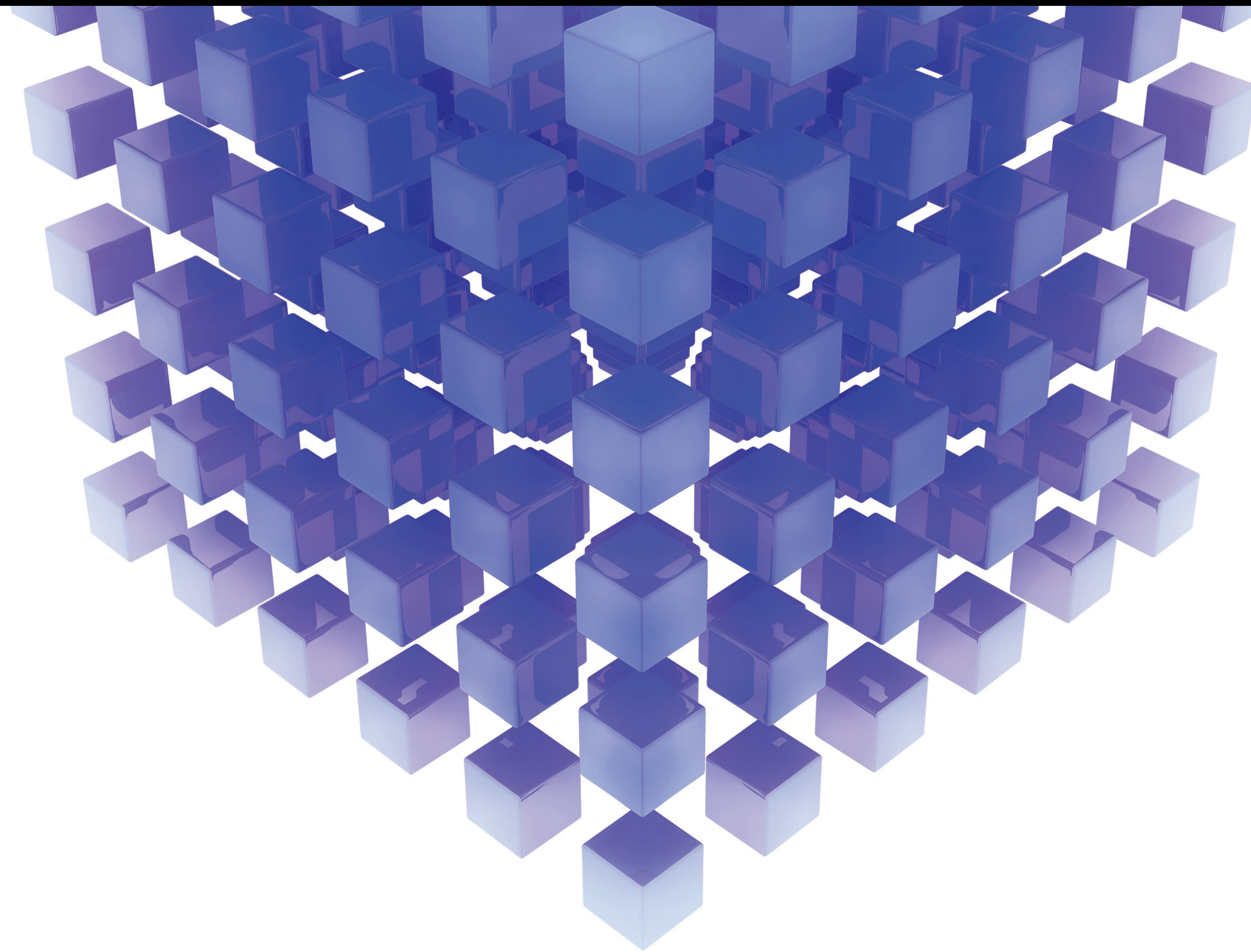
Mathematical Problems in Engineering

# Bilevel Programming, Equilibrium, and Combinatorial Problems with Applications to Engineering

Guest Editors: Vyacheslav Kalashnikov, Timothy I. Matis, José Fernando Camacho Vallejo, and Sergii V. Kavun



# Bilevel Programming, Equilibrium, and Combinatorial Problems with Applications to Engineering

# Bilevel Programming, Equilibrium, and Combinatorial Problems with Applications to Engineering

Guest Editors: Vyacheslav Kalashnikov, Timothy I. Matis, José Fernando Camacho Vallejo, and Sergii V. Kavun

# Editorial Board

# Contents

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 490758, 3 pages
http://dx.doi.org/10.1155/2015/490758

*Editorial*

# Bilevel Programming, Equilibrium, and Combinatorial Problems with Applications to Engineering

**Vyacheslav Kalashnikov,[1,2,3] Timothy I. Matis,[4] José Fernando Camacho Vallejo,[5] and Sergii V. Kavun[6]**

[1] *Department of Systems & Industrial Engineering, Tecnológico de Monterrey (ITESM), Campus Monterrey, Avenida Eugenio Garza Sada 2501 Sur, 64849 Monterrey, NL, Mexico*
[2] *Department of Social Modeling, Central Economics & Mathematics Institute (CEMI), Russian Academy of Sciences, Nakhimovsky pr. 47, Moscow 117418, Russia*
[3] *Department of Computing, Sumy State University, Rimsky-Korsakov St. 2, Sumy 40007, Ukraine*
[4] *Department of Industrial Engineering, Texas Tech University, P.O. Box 43061, Lubbock, TX 79409, USA*
[5] *Facultad de Ciencias Físico-Matemáticas, Universidad Autónoma de Nuevo León, Avenida Universidad S/N, 66450 San Nicolás de los Garza, NL, Mexico*
[6] *Kharkiv Institute of Banking of the University of Banking of the National Bank of Ukraine, Peremogy Avenue 55, Kharkiv 61174, Ukraine*

Correspondence should be addressed to Vyacheslav Kalashnikov; kalash@itesm.mx

Received 16 September 2014; Accepted 16 September 2014



Although a wide range of applications fit the bilevel programming framework, real-life implementations are scarce, due mainly to the lack of efficient algorithms for tackling medium- and large-scale bilevel programming problems. Solving a bilevel (more generally, hierarchical) optimization problem, even in its simplest form, is a difficult task. A lot of different alternative methods may be used based on the structure of the problem analyzed, but there is no general method that guarantees convergence, performance, or optimality for every type of problem.

Many new ideas appeared and were discussed in works of plenty of authors. Among them, we would name Dempe [1], Dempe et al. [2], Dempe and Dutta [3], Dewez et al. [4], Thi et al. [5], and Vicente and Calamai [6], whose works have developed various ways of reducing original bilevel programming problems to equivalent single level ones, thus making their solution somewhat an easier task for conventional mathematical programming software packages.

Mixed-integer bilevel programming problems (with part of variables at the upper and/or lower level being integer/Boolean ones) are even harder for the well-known conventional optimization techniques. For instance, a usual replacement of the lower level optimization problem with the corresponding Karush-Kuhn-Tucker (KKT) conditions may not work if some lower level variables are not continuous. Therefore, solid theoretical base is necessary to be found, in order to propose efficient algorithmic procedures aimed at finding local or global solutions of such a problem.

A great amount of new applied problems in the area of energy networks have recently arisen that can be efficiently solved only as mixed-integer bilevel programs. Among them are the natural gas cash-out problem, the deregulated electricity market equilibrium problem, biofuel problems, problem of designing coupled energy carrier networks, and so forth, if we mention only part of such applications. Bilevel models to describe migration processes are also in the running of the most popular new themes in the area of bilevel programming.

This special volume comprises papers dealing with three main themes: bilevel programming, equilibrium models, and combinatorial (integer programming) problems and their applications to engineering.

The special issue opens with a survey paper "*Bilevel programming and applications*" by S. Dempe et al. that sums up some recent and new directions and results of the development of the mathematical methods aimed at the solution of bilevel programs of different types and their applications to real-life problems.

One of such applications is the well-known bilevel programming problem arising in the natural gas industry and related to the solution of the important imbalance cash-out problem (a good part of the above-mentioned survey paper is devoted to this area). Since the stochastic statement of the problem in question needs the generation of scenario trees with a very large number of branches and arcs, the reduction of the quantity of common variables is an important challenge. This item is analyzed and solved by proposing new methods of grouping the variables (classifying and representing various groups of states with respect to the natural gas prices) in the article "*US natural gas market classification using pooled regression*" by V. V. Kalashnikov et al. The techniques proposed in this paper help classify regions or states into groups or classes that share similar regression parameters. Once obtained, these groups may be used to make assumptions about corresponding natural gas prices in further studies.

One of the principal management problems dealt with in numerous organizations and institutions in the public sector is to decide how to invest and manage funds available to support potential research and development projects in various areas. A classification of the portfolio problem distinguishes two types: (a) static and (b) dynamic. In (a) only projects proposals for funding are involved, while in (b) at certain moments some active projects are withdrawn from the portfolio and some inactive projects are activated. Within this framework, a *social project* is a group of tasks or activities consuming funds, carried out during a given period of time in one or more regions, with an impact on the objectives set by an organization, and focused on providing solution to problems or needs of the society. The paper "*Selecting large portfolios of social projects in public organizations*" by I. Litvinchev et al. models the task as a two-objective mixed-integer linear programming problem. The model supports both complete (all or nothing) and partial (a certain amount from a given interval of funding) resource allocation policies. Numerical results for large scale problem instances are also presented. The computational experiment demonstrates that increasing the number of projects subject to funds reduction results in a decrease in both objectives. However, if the number of projects subject to funds reduction is fixed but the strength of funds reduction is increased, one of the objectives may grow at the expense of a larger drop in another.

A capacitated fixed-cost facility location problem with transportation choices (CFCLP-TC) is studied in the paper "*Variations in the flow approach to CFCLP-TC for multi-objective supply chain design*" by M. P. Hertwin et al. The problem based on a production network of two echelons with multiple plants, a set of potential distribution centers and customers, is formulated as an optimization model with two objective functions with time and cost as variables. The main contribution of this paper is in an extension of the existing approaches to the supply of products to customers through multiple sources, the direct flow between plants and customers, and the product flow between distribution centers. Based upon these approaches, the authors generate mathematical programming models and propose solution methods by the $\varepsilon$-constraint approach. Namely, they generate Pareto boundaries and thus compare each of those approaches with the original model. The models are implemented in GAMS and solved with CPLEX.

Another work presenting a novel capacitated model for supply chain network design is the paper "*Design and optimization of capacitated supply chain networks including quality measures*" by K. K. Castillo-Villar et al. The model in question (abbreviated as SCND-COQ) is described as a mixed-integer nonlinear programming problem, which can be used at a strategic planning level to design a supply chain network that maximizes the total profit subject to meeting an overall quality level of the final product at the minimum costs. To process the problem, the authors propose five combinatorial optimization algorithms based on nonlinear optimization, heuristic, and metaheuristic approaches, which are used to solve realistic instances of practical size. The computational results demonstrate that the state and individual representation in the simulated annealing and the genetic algorithm, respectively, has a significant impact on the improvement of the solution quality.

A problem of packing a limited number of unequal circles in a fixed size rectangular container is considered in the paper "*Integer programming formulations for approximate packing circles in a rectangular container*" by I. Litvinchev and E. L. O. Espinosa. The aim of this research is to maximize the (weighted) number of circles placed into the container or minimize the waste. This problem has numerous applications in logistics, including production and packing for the textile, apparel, naval, automobile, aerospace, and food industries. Frequently the problem is formulated as a nonconvex continuous optimization problem, which is solved by heuristic techniques combined with local search procedures. The authors propose new formulations for an approximate solution of the packing problem. The container is approximated by a regular grid and the nodes of the grid are considered as potential positions for assigning centers of the circles. The packing problem is then stated as a large scale linear 0-1 optimization problem. The binary variables represent the assignment of centers to the nodes of the grid. Nesting circles inside one another are also considered. The resulting binary problem is then solved by commercial software. Numerical results are presented to demonstrate the efficiency of the proposed approach and compared with known results.

The paper "*A heuristic procedure for a ship routing and scheduling problem with variable speed and discretized time windows*" by K. K. Castillo-Villar et al. develops a heuristic algorithm solving a routing and scheduling problem for tramp shipping with discretized time windows. The problem consists of determining the set of cargoes that should be served by each ship, as well as the arrival, departure, and waiting times at each port, while minimizing the total costs. The heuristic proposed is based on a variable neighborhood search, considering a number of neighborhood structures to find a solution to the problem. The authors present computational results, and, for comparison purposes, they consider instances that can be solved directly by CPLEX to test the performance of the proposed heuristic. The heuristic achieves good solution qualities within reasonable computational times. The obtained numerical results are encouraging and make the presented heuristic quite promising to be useful when solving large real-size examples.

It is well-known that modern approaches to the development of a national economy are often characterized with an imbalanced inflation of some economic branches leading to a disproportional socioeconomic territories development (SETD). Such disproportions, together with other similar factors, frequently result in a lack of economic integrity, various regional crises, and a low rate of the economic and territorial growth. Those disproportions may also conduce to an inadequate degree of the interregional collaboration. The paper "*Simulation of territorial development based on fiscal policy tools*" by R. Brumnik et al. proposes some new ways of regulating imbalances in the territorial development making use of the fiscal policy tools. The latter can immediately reduce the amplitude of economic cycle fluctuations and provide a stable development of the economic state system. The same approach is applied to control the processes of transformation of the tax legislation and tax relations, as well as the levying and redistribution of the recollected taxes among the territories' budgets (this approach is also known as a tax policy). To resume, this paper describes comprehensive models of financial regulation of the socioeconomic territorial development that can help in estimating and choosing the right financial policy parameters. The research highlights the following conclusions: analysis of the predictive dynamics of socioeconomic development territories, in the case of the implementation of an optimistic scenario of tax revenue, demonstrates the effectiveness of the adopted stabilization policy.

Though a bit off the special issue's general stream, still interesting results of research are provided by the article "*The solution of fourth order boundary value problem arising out of the beam-column theory using Adomian decomposition method*" by O. Kelesoglu. In this study, Adomian decomposition method (ADM) is applied to a linear nonhomogeneous boundary value problem arising from the beam-column theory. The obtained results of numerical experiments are presented in tables and illustrated with graphs. The proposed approach provides a rapidly converging method, with the approximate solutions tending to the exact solution of the problem in question. This fact characterizes the method as appropriate and reliable for such kind of problems.

We hope that readers of this special issue will find not only new ideas and algorithms dealing with the difficult applied problems like supply chain design, assignment problems, bilevel programming models, and so forth, but also some interesting results related to regression analysis, as well as the modern simulation techniques.

## Acknowledgments

A special gratitude is also due to the long list of referees, whose valuable comments and suggestions have helped us a lot in improving our results and/or presentation style. The research activity of Vyacheslav Kalashnikov was financially supported by Tecnológico de Monterrey (ITESM), Campus Monterrey, and by the SEP-CONACYT Project CB-2013-01-221676, Mexico. The research activity of José Fernando Camacho Vallejo was financially supported by the Project PROMEP/103.5/10/3889, the Academic Groups Research Project PROMEP/103.5/12/4953, and the Autonomous University of Nuevo Leon (UANL) within the Support Program for Scientific Research and Technology (PAICYT) within the Project CE960-11.

*Vyacheslav Kalashnikov*
*Timothy I. Matis*
*José Fernando Camacho Vallejo*
*Sergii V. Kavun*

## References

[1] S. Dempe, *Foundations of Bilevel Programming*, Springer, Heidelberg, Germany, 2002.

[2] S. Dempe, B. S. Mordukhovich, and A. B. Zemkoho, "Necessary optimality conditions in pessimistic bilevel programming," *Optimization*, vol. 63, no. 4, pp. 505–533, 2014.

[3] S. Dempe and J. Dutta, "Is bilevel programming a special case of a mathematical program with complementarity constraints?" *Mathematical Programming A*, vol. 131, pp. 37–48, 2012.

[4] S. Dewez, M. Labbé, P. Marcotte, and G. Savard, "New formulations and valid inequalities for a bilevel pricing problem," *Operations Research Letters*, vol. 36, no. 2, pp. 141–149, 2008.

[5] L. Thi, T. Duc, and P. Dinh, "A DC programming approach for a class of bilevel programming problems and its application in portfolio selection," *Numerical Algebra, Control and Optimization*, vol. 2, no. 1, pp. 167–185, 2012.

[6] L. N. Vicente and P. H. Calamai, "Bilevel and multilevel programming: a bibliography review," *Journal of Global Optimization*, vol. 5, no. 3, pp. 291–306, 1994.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 310301, 16 pages
http://dx.doi.org/10.1155/2015/310301

*Review Article*

# Bilevel Programming and Applications

**Vyacheslav V. Kalashnikov,[1,2,3] Stephan Dempe,[4] Gerardo A. Pérez-Valdés,[5] Nataliya I. Kalashnykova,[6] and José-Fernando Camacho-Vallejo[6]**

[1] *Departamento de Ingeniería Industrial y de Sistemas, Tecnológico de Monterrey, Avenida Eugenio Garza Sada 2501 Sur, 64849 Monterrey, NL, Mexico*
[2] *Department of Social Modeling, Central Economics and Mathematics Institute (CEMI) of the Russian Academy of Sciences (RAS), Nakhimovsky Prospekt 17, Moscow 117418, Russia*
[3] *Department of Electronics and Computing, Sumy State University, Rimsky-Korsakov Street 2, Sumy 40007, Ukraine*
[4] *Institut für Numerische Mathematik und Optimierung, Fakultät für Mathematik und Informatik, TU Bergakademie Freiberg, Akademiestraße 6, 09596 Freiberg, Germany*
[5] *SINTEF, Box 4760, Sluppen, 7465 Trondheim, Norway*
[6] *Facultad de Ciencias Físico-Matemáticas (FCFM), Universidad Autónoma de Nuevo León (UANL), Avenida Universidad S/N, 66450 San Nicolás de los Garza, NL, Mexico*

Correspondence should be addressed to Vyacheslav V. Kalashnikov; kalash@itesm.mx

Received 25 June 2014; Accepted 23 July 2014

Academic Editor: Sergii V. Kavun



A great amount of new applied problems in the area of energy networks has recently arisen that can be efficiently solved only as mixed-integer bilevel programs. Among them are the natural gas cash-out problem, the deregulated electricity market equilibrium problem, biofuel problems, a problem of designing coupled energy carrier networks, and so forth, if we mention only part of such applications. Bilevel models to describe migration processes are also in the list of the most popular new themes of bilevel programming, as well as allocation, information protection, and cybersecurity problems. This survey provides a comprehensive review of some of the above-mentioned new areas including both theoretical and applied results.

## 1. Introduction

Although a wide range of applications fit the bilevel programming framework, real-life implementations are scarce, due mainly to the lack of efficient algorithms for tackling medium- and large-scale bilevel programming problems (BLP). Solving a bilevel (more generally, hierarchical) optimization problem, even in its simplest form, is a difficult task. A lot of different alternative methods may be used based on the structure of the problem analyzed, but there is no general method that guarantees convergence, performance, or optimality for every type of problem.

Many new ideas appeared and were discussed in works of plenty of authors. Among them, we would name Dempe [1], Dempe et al. [2], Dempe and Dutta [3], Dewez et al. [4], Thi et al. [5], and Vicente and Calamai [6], whose works have developed various ways of reducing original bilevel programming problems to equivalent single-level ones, thus making their solution somewhat easier task for conventional mathematical programming software packages.

Mixed-integer bilevel programming problems (with part of the variables at the upper and/or lower level being integer/Boolean ones) are even harder for the conventional optimization techniques. For instance, a usual replacement of the lower level optimization problem with a corresponding KKT condition may not work if some of the lower level variables are not continuous. Therefore, solid theoretical base is necessary to be found, in order to propose efficient algorithmic procedures aimed at finding local or global solutions of such a problem.

A great amount of new applied problems in the area of energy networks has recently arisen that can be efficiently solved only as mixed-integer bilevel programs. Among them are the natural gas cash-out problem, the deregulated electricity market equilibrium problem, biofuel problems, a problem of designing coupled energy carrier networks, and so forth, if we mention only part of such applications. Bilevel models to describe migration processes are also in the list of the most popular new themes of bilevel programming.

This special volume of the Hindawi journal Mathematical Problems in Engineering comprises papers dealing with three main themes: bilevel programming, equilibrium models, and combinatorial (integer programming) problems, and their applications to engineering. Because of that, it opens with this survey paper "Bilevel Programming and Applications" summing up some recent and new directions and results of the development of the mathematical methods aimed at the solution of bilevel programs of different types and their applications to real-life problems.

The paper is organized as follows: the survey of the literature dealing with the formulation and history of bilevel programming problems is given in Section 2. Section 3 describes the ways the linear bilevel programs are treated, while Section 4 surveys the recent results in an important application of BLP to the well-known imbalance cash-out problem arising in the natural gas industry. Section 5 reviews the new methods of reducing the number of upper level variables, which helps a lot in applying stochastic programming algorithms to solve the optimal cash-out problems. Section 6 describes various promising bilevel approaches to the mixed-integer allocation model. Finally, Section 7 presents the latest bilevel mechanisms to solve very important information protection and cybersecurity problems. The conclusion, acknowledgements, and the list of references finish the survey.

## 2. Bilevel Programs: Statement and History

A *bilevel program* is an optimization problem where the feasible set is partly determined through a solution set mapping of a second parametric optimization problem [1]. The latter problem is given as

$$\min_{y} \{f(x,y) : g(x,y) \le 0, y \in T\}, \qquad (1)$$

where $f : R^n \times R^m \to R, g : R^n \times R^m \to R^p, T \subseteq R^m$ is a (closed) set.

Let $Y : R^n \to R^m$ denote the feasible set mapping: let

$$Y(x) := \{y : g(x,y) \le 0\},$$
$$\varphi(x) := \min_{y} \{f(x,y) : g(x,y) \le 0, y \in T\} \qquad (2)$$

be the optimal value function, and let $\Psi : R^n \to R^m$ be the solution set mapping of the problem (1) for a fixed value of $x$:

$$\Psi(x) := \{y \in Y(x) \cap T : f(x,y) \le \varphi(x)\}. \qquad (3)$$

Let

$$\mathbf{gph}\Psi := \{(x,y) \in R^n \times R^m : y \in \Psi(x)\} \qquad (4)$$

be the graph of the mapping $\Psi$. Then, the bilevel programming problem is given as

$$\text{“}\min_{x}\text{”} \{F(x,y) : G(x) \le 0, (x,y) \in \mathbf{gph}\Psi, x \in X\}, \qquad (5)$$

where $F : R^n \times R^m \to R, G : R^n \to R^q, X \subseteq R^n$ is a closed set.

Problems (1) and (5) can be interpreted as an hierarchical game of two decision makers (or players) which make their decisions according to the hierarchical order. The first player (called the *leader*) makes his selection first and communicates it to the second player (the so-called *follower*). Then, knowing the choice of the leader, the follower selects his *response* as an optimal solution of problem (1) and gives this back to the leader. Thus, the leader's task is to determine a best decision, that is, a point $\hat{x}$ which is feasible for problem (5): $G(\hat{x}) \le 0, \hat{x} \in X$, minimizing, together with the response $\hat{y} \in \Psi(\hat{x})$, the function $F(x,y)$. Therefore, problem (1) is called the follower's problem and (5) the leader's problem. Problem (5) is the *bilevel programming problem*.

*2.1. Optimistic and Pessimistic Approaches.* Strictly speaking, problem (5) is ill-posed in the case when the set $\Psi(x)$ is not a singleton for some $x$, which means that the mapping $x \mapsto F(x, y(x))$ is not a function but a multivalued mapping. This is implied by an ambiguity in the computation of the upper level objective function value, which is rather an element in the set $\{F(x,y) : y \in \Psi(x)\}$. The quotation marks in (5) are used purely to indicate this ambiguity. To cope with such an obstacle, there are several ways out.

(1) The leader can assume that the follower is willing (and able) to cooperate. In this case, the leader simply selects the solution within the set $\Psi(x)$ that is the best one with respect to the upper level objective function. This leads then to the function

$$\varphi_o(x) := \min \{F(x,y) : y \in \Psi(x)\} \qquad (6)$$

to be minimized over the set $\{x : G(x) \le 0, x \in X\}$. This is the *optimistic* approach leading to the *optimistic bilevel programming problem*. Roughly speaking, this problem is closely related to the problem

$$\min_{x,y} \{F(x,y) : G(x) \le 0, (x,y) \in \mathbf{gph}\Psi, x \in X\}. \qquad (7)$$

If $\overline{x}$ is a local minimum point of the function $\varphi_o(\cdot)$ on the set

$$\{x : G(x) \le 0, x \in X\} \qquad (8)$$

and $\overline{y} \in \Psi(\overline{x})$, then the point $(\overline{x}, \overline{y})$ is also a local minimum point of problem (7). The converse is in general not true. For more information about the relation between both problems, the interested reader is referred to Dempe [1].

(2) The leader has no possibility to influence the follower's selection, neither has he/she any guess about the follower's choice. In this case, the leader has to take into account the follower's ability to select the worst solution with respect to the leader's objective function; hence the leader has to diminish the damage resulting from such an unlucky selection. This brings up the function

$$\varphi_p(x) := \max \{F(x, y) : y \in \Psi(x)\} \tag{9}$$

to be minimized on the set $\{x : G(x) \leq 0, x \in X\}$ :

$$\min \left\{\varphi_p(x) : G(x) \leq 0, x \in X\right\}. \tag{10}$$

This is the *pessimistic* approach resulting in the *pessimistic bilevel programming problem*. This problem is often much more complicated than the optimistic bilevel programming problem; see Dempe [1].

There is also another pessimistic bilevel optimization problem in the literature. To describe this problem consider the bilevel optimization problem with connecting upper level constraints and an upper level objective function depending only on the upper level variable $x$:

$$\text{“}\min_x\text{”} \{F(x) : G(x, y) \leq 0, y \in \Psi(x)\}. \tag{11}$$

In this case, a point $x$ is feasible if there exists $y \in \Psi(x)$ such that $G(x, y) \leq 0$, which can be written as

$$\min_x \{F(x) : G(x, y) \leq 0 \text{ for some } y \in \Psi(x)\}. \tag{12}$$

Now, if the quantifier $\exists$ is replaced by $\forall$ we derive a second pessimistic bilevel programming problem

$$\min_x \{F(x) : G(x, y) \leq 0 \ \forall y \in \Psi(x)\}. \tag{13}$$

This problem has been investigated in Wiesemann et al. [7]. The relations between (13) and (10) should be studied in the future.

(3) The leader is able to predict a selection of the follower: $y(x) \in \Psi(x)$ for all $x$. If this function is inserted into the upper level objective function, this leads to the problem

$$\min_x \{F(x, y(x)) : G(x) \leq 0, x \in X\}. \tag{14}$$

Such a function $y(\cdot)$ is called a selection function of the point-to-set mapping $\Psi(\cdot)$. Hence, we call this approach the *selection function approach*. One special case of this approach arises if the optimal solution of the lower level problem is unique for all values of $x$. It is obvious that the optimistic and the pessimistic problems are special cases of the selection function approach.

Even under restrictive assumptions (as in the case of linear bilevel optimization or if the follower's problem has a unique optimal solution for all $x$), the function $y(\cdot)$ is in general nondifferentiable. Hence, the bilevel programming problem is a *nonsmooth* optimization problem.

Various results and examples/counterexamples concerning the existence of solutions to different formulations of bilevel programming problems can be found in [1, 8–10], to mention only few.

*2.2. A Short History of Bilevel Programming.* The history of bilevel programming dates back to von Stackelberg who (in 1934 in monograph [11]) formulated a hierarchical game of two players now called *Stackelberg game*. The formulation of the bilevel programming problem goes back to Bracken and McGill [12]; the notion “bilevel programming” has been coined probably by Candler and Norton [13]; see also Vicente et al. [14]. With the beginning of the 80's of the last century a very intensive investigation of bilevel programming started. A number of monographs, for example, Bard [15], Shimizu et al. [16] and Dempe [1], edited volumes, see Dempe and Kalashnikov [17] and Migdalas et al. [18] and (annotated) bibliographies, for example, Vicente and Calamai [6], and Dempe [19] have been published in that field.

One possibility to investigate bilevel programs is to transform them into single-level (or ordinary) optimization problems. In the first years, linear bilevel programming problems (where all the involved functions are affine (linear) and the sets $X$ and $T$ are whole spaces) were usually transformed making use of linear programming duality or, equivalently, the Karush-Kuhn-Tucker conditions for linear programming. Applying this approach, solution algorithms have been developed; *compare*, for example, Candler and Townsley [20]. The transformed problem is a special case of a mathematical program with equilibrium constraints MPEC (now sometimes called *mathematical program with complementarity constraints*, MPCC). We can call this the *KKT transformation* of the bilevel programming problem. This approach is also possible for convex parametric lower level problems satisfying some regularity assumption.

General MPCCs have been the topic of some monographs; *see* Luo et al. [21] and Outrata et al. [22]. Solution algorithms for MPCCs (see, for instance, Outrata et al. [22], Demiguel et al. [23], Leyffer et al. [24], and many others) have also been suggested for solving bilevel programming problems.

Since MPCCs are nonconvex optimization problems, solution algorithms will hopefully compute local optimal solutions of the MPCCs. Thus, it is interesting if a local optimal solution of the KKT transformation of a bilevel programming problem is related to a local optimal solution of the latter problem. This has been the topic of the paper by Dempe and Dutta [3].

Later on, the selection function approach to bilevel programming has been investigated in the case when the optimal solution of the lower level problem is uniquely determined and strongly stable in the sense of Kojima [25]. Then, under some assumptions, the optimal solution of the lower level problem is a $PC^1$-function; see Ralph and Dempe [26] and Scholtes [27] for the definition and properties of $PC^1$-functions. This can then be used to determine necessary and sufficient optimality conditions for bilevel programming; *compare* Dempe [28].

Using the optimal value function $\varphi(x)$ of the lower level problem (1), the bilevel programming problem (7) can be replaced with

$$\min_{x,y} \{F(x,y) : G(x) \leq 0, g(x,y) \leq 0, \\ f(x,y) \leq \varphi(x), x \in X\}. \tag{15}$$

This is the so-called *optimal value transformation*.

Since the optimal value function is nonsmooth even under very restrictive assumptions, this is a nonsmooth, nonconvex optimization problem. Using nonsmooth analysis, see, for example, Mordukhovich [29, 30] and Rockafellar and Wets [31], optimality conditions for the optimal value transformation can be obtained, *compare* Outrata [32], Ye and Zhu [33], and Dempe et al. [34].

Nowadays, a large number of Ph.D. theses have been written on bilevel programming problems, very different types of (necessary and sufficient) optimality conditions can be found in the literature, the number of applications is huge, and both exact and heuristic solution algorithms have been suggested.

## 3. Linear Bilevel Programming Problems

The linear bilevel program is the problem of the following structure:

$$\min_{x,y} \{a^\top x + b^\top y : Ax + By \leq c, (x,y) \in \mathbf{gph}\Psi\}, \tag{16}$$

where $\Psi(\cdot)$ is the solution set mapping of the lower level problem

$$\Psi(x) := \operatorname{Arg}\min_{y} \{d^\top y : Cy \leq x\}. \tag{17}$$

Here, $A$, $B$, and $C$ are matrices of sizes $p \times n$, $p \times m$, and $n \times m$, respectively, and all variables and vectors used are of appropriate dimensions. Note that we have used here the so-called optimistic bilevel optimization problem, which is related to problem (7).

The so-called *connecting constraints* $Ax + By \leq c$ are included in the upper level problem. Validity of such constraints is beyond the selection of the leader and can be verified only after the follower has selected his/her (possibly *not* unique) optimal solution. In the case especially when $\Psi(x)$ does not reduce to a singleton, certain difficulties may arise. In order to examine the bilevel programming problem in the case that $\Psi(x)$ does not reduce to a singleton, Ishizuka and Aiyoshi [35] introduced their double penalty method. In general, connecting constraints may imply that the feasible set of the bilevel programming problem is disconnected. This situation is illustrated by the following example:

*Example 1* (Mersha and Dempe [36]). Consider the problem

$$z = -x - 2y \longrightarrow \min_{x,y} \tag{18}$$

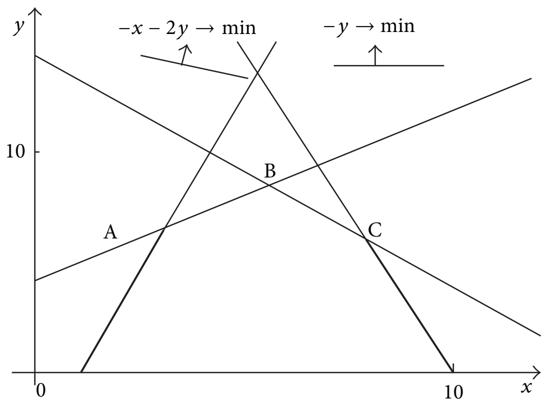



Figure 1: The problem with upper level connecting constraints. The feasible set is depicted with bold lines. The point $C$ is global optimal solution and point $A$ is a local optimal solution.

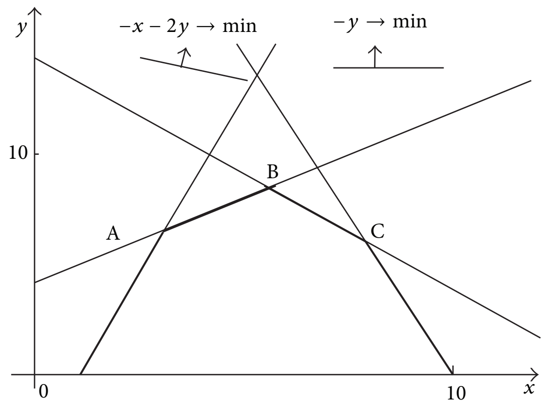



Figure 2: The problem when the upper level connecting constraints are shifted into the lower level problem. The feasible set is depicted with bold lines. The global optimal solution is point $B$.

subject to

$$2x - 3y \geq -12 \\ x + y \leq 14, \\ y \in \operatorname{Arg}\min_{y} \{-y : -3x + y \leq -3, 3x + y \leq 30\}. \tag{19}$$

The optimal solution for this problem is the point $C$ at $(\overline{x}, \overline{y}) = (8, 6)$ (see Figure 1). But if we shift the two upper level constraints to the lower level we get the point $B$ at $(\tilde{x}, \tilde{y}) = (6, 8)$ as an optimal solution (see Figure 2). From this example one can easily notice that if we shift constraints from the upper level to the lower one, the optimal solution obtained prior to shifting is not optimal any more in general. Hence ideas based on shifting constraints from one level to another will lead to a solution which may not solve the original problem.

In Example 1, the optimal solution of the lower level problem was unique for all $x$. If this is not the case, feasibility of a selection of the upper level decision maker possibly depends on the selection of the follower. In the optimistic case, this means that the leader selects within the set of optimal solutions of the follower's problem one point which is at the same time feasible for the upper level connecting constraints and gives the best objective function value for the upper level objective function.

As we can see in Example 1 the existence of connecting upper level constraints might lead, in general, to a *disconnected* feasible set in the bilevel programming problem. Therefore, solution algorithms will live in one of the connected components of the feasible set (i.e., a sequence of feasible points which all belong to one of the connected parts is computed) or they need to jump from one of the connected parts of the feasible set to another one. Some ideas of discrete optimization are needed in such cases.

In order to avoid the above-mentioned difficulties, some researchers restrict themselves to the cases when the upper level constraints depend on the upper level variables only (i.e., matrix $B$ is zero-matrix, $B = 0$). Thus, the bilevel problem (16)-(17) reduces to a simpler one:

$$\min_{x,y} \left\{ a^\top x + b^\top y : Ax \le c, (x, y) \in \mathbf{gph}\Psi \right\}, \quad (20)$$

where $\Psi(\cdot)$ is the solution set mapping of the lower level problem

$$\Psi(x) := \operatorname{Arg}\min_{y} \left\{ d^\top y : Cy \le x \right\}. \quad (21)$$

In this problem, parametric linear optimization (see, for example, Nožička et al. [37]) can be used to show that the graph of the mapping $\Psi(\cdot)$ equals the connected union of faces of the set $\{(x, y)^\top : Cy \le x\}$.

## 4. Application of Bilevel Programming to Imbalance Cash-Out Problem

In the early 1990s, several regulations were passed in the USA and the European Union [38, 39] changing the way natural gas was marketed and traded. Particularly, this liberalization [40] effectively ended a period in which natural gas was a state-driven industry. The liberalization has also created the emergent natural gas markets, as well as a strong demand for models to better tackle the new problems and profit from this new setting [41, 42]. It is possible to say that the above-mentioned processes formed the natural gas supply chain. The resulting market configuration demanded the independence of the transportation and commercialization processes. As a result of this paradigm shift—and the accompanying restructurization of the market—a systematic analysis of several new features becomes indispensable.

Of particular interest is a problem that arises in the natural gas supply chain, namely, that of *balancing* the fuel volumes over a distribution network. Such a balancing procedure directly concerns the Pipeline Operating Company (POC), since the correct operation of the pipeline means the well controlled volumes of the transported gas. Moreover, any natural gas shipping company (NGSC) is also concerned with the balancing of the volumes because it is often impossible to avoid an *imbalance* justified by certain economic reasons. A natural gas shipping company's business is to sell the gas by moving it through the pipeline to its clients: it has to fulfill signed contracts first and then market excesses of the gas to achieve the maximum profits. In order to do that, the NGSC has to manage the volumes at each selling point (so-called *pipeline meters*) taking into account the balance, the selling prices, and the total revenue. The basic mathematical framework of this problem's modeling is found in [43].

Owing not only to this liberalization, but also to the new local conditions that are more open to competition, new small players entered the natural gas industry, especially at the local scale. Indeed, the USA has over 80 interstate, long-distance pipelines [44], serving different regions with various climatic, demographic, economic, and political circumstances. Natural gas usage in Alabama, for example, intuitively is not the same as in Oregon; thus the market dynamics of the fuel are also different, and this, we presume, should be reflected in some way in the econometric data of the states.

Not only macroeconomic trends, however, are affected by this setting. When doing cross-regions studies of various aspects of the supply chain, such as the forecasting of demand [45, 46], the balancing of the pipelines after imbalances have been created by the natural gas shippers [43, 47, 48], or the dynamics of interstate-intrastate systems [49], one has to take into account the existence of different markets.

The existence of a common relationship between price and consumption of natural gas across several zones allows for strong claims of uniformity, which are useful when, for example, we are building scenarios for a stochastic problem. Indeed, if we manage to group the regions in clusters with similar price and consumption functions, we can reduce the number of variables needed in a scenario tree formulation [42, 50]

It must be emphasized that, while natural gas pipeline networks have been thoroughly studied, most of the existent models focus on aspects of this part of the supply chain other than the NGSC-POC interaction in the system balancing, such as network operation optimization [51, 52] or deployment of facilities [53]. There are also papers considering the natural gas supply chain in a multilevel scheme, in which both the NGSC and the POC are present and accounted for, such as the related [54, 55]. These works are remarkable in the sense that they span the whole supply chain with much emphasis on the traders (financial front-ends of the natural gas producers,) so that there is little to no mention of imbalances in the system resulting from the dealings of the NGSCs and the POCs, even though both actors are present in the models.

Many authors do acknowledge [56, 57] the existence of a problematic situation in the NGSC-POC system following the paradigm shift, yet we have found very few sources that explain plausible ways in which this problem is nowadays solved. For example, [58] shows how storage is required by the NGSC when no flexibility exists in the network volume management, either because it is not allowed, or because it is not technically possible. Nevertheless, balancing is an important part of the modern natural gas supply chain management, and to date, no policy has been accepted as optimal regarding the way, in which the imbalances produced by the NGSC are physically and economically handled. Among the most important tools that aid the POC in its task of restoring the balance of the system is the *arbitrage penalization policies*, in which the POC performs a maintenance redistribution of the imbalances in the system and charges the NGSC(s) for the cost of this operation.

In [59], one finds a modeling framework (which we are going to follow) of the penalization part of this problem. This penalization refers only to the cash-out that occurs between the NGSC and the POC: it leaves outside any reference to actual market conditions, which are obviously important to the NGSC. The paper presents a solution method through a modification of the original problem, as well as the analysis of how this modification affects the objective function and the obtained solutions. In [47], the authors compare two algorithms that solve the problem making use of certain numerical procedures. In the present paper, we adapt these algorithms to our extended model. We also make use of the idea proposed in [43] to divide our problem into several *generalized transportation problems* when finding its numerical solution.

In [60, 61], we study a modified version of the above-described problem, in which the upper level objective function includes certain new terms based upon the net profit of the leader—the natural gas shipping company. This formulation assumes, however, the complete knowledge (perfect information) about the changes in the prices of natural gas during the process, which is somewhat nonrealistic and not quite useful, as the resulting function does not clearly reflect the reasons behind the actions of the NGSC.

In [62] a stochastic reformulation of the problem is presented, so that the NGSC is now able to forecast the next several values of the natural gas demand and then to plan the extraction of natural gas from the pipeline. The resulting model is a stochastic variation of the original mixed-integer bilevel optimization problem, for which two different solution methods are proposed and compared.

To the best of our knowledge, there is no literature, beyond the works listed in the paragraphs above and their derivatives, that explicitly deals with the NGSC-POC subsystem in the same way we propose, formulating a bilevel optimization problem out of the balancing operations. We attribute this to the relatively recent nature of the problem we are dealing with, as well as the difficulty of its accurate formulation for specific instances.

*4.1. Problem Specification.* Following the scheme constructed in [47, 59], we will consider a leader-follower system, in which the first agent (the leader), namely, the Natural Gas Shipping Company (NGSC), buys the gas at the wells, arranges for its injection into an (interstate) pipeline at its starting point, and extracts some amount of gas—ideally, equal to the deposited amount—from pipeline meters in several *pool zones* across the country. The follower here is the administrator of the pipeline, which we call the Pipeline Operating Company (POC), who permits the NGSC to extract amounts of natural gas that may differ from the originally injected volumes, thus creating positive or negative imbalances. The latter is a kind of usual market practice that allows for a dynamic flow of the fuel within the natural gas supply chain.

However, since disrupting the system in this way implies extra costs for the NGSC, the company attempts to do it only when its predictions of future market conditions show that the total revenues overcome the losses incurred by the penalization policy applied to the NGSC. It is clear that the NGSC needs tools that provide it with the best possible information and help it make advantage of the latter.

The NGSC-POC system operates in the following way.

(1) The NGSC makes a forecast of the demand it is likely to have during the next period (month, year, etc.) and considers different scenarios, in which this can occur.

(2) The NGSC books certain capacity $D^c$ for every pool zone and stage (day, week, month, etc.)

(3) For each subsequent stage, the NGSC determines the amount of gas to extract and sell, which possibly creates positive and negative imbalances in the process; this continues until the period is over;.

(4) The POC studies the resulting last day imbalances and rearranges them according to certain business rules.

(5) The POC charges the NGSC with certain penalty for the final (rearranged) imbalances. The latter may occur to be negative; that is, the POC may pay to the NGSC.

(6) The NGSC calculates the net profits as its sales revenue minus the penalty.

The resulting model is a bilevel multistage stochastic optimization problem [63], in which the upper level decision maker (the leader) is the NGSC who has the objective of maximizing its net profit as the revenue from the sales of its gas in the pipeline minus the penalty imposed by the POC. The lower level decision maker (the follower) is the POC who aims at minimizing the absolute value of the penalization cash-out flow, either from the POC to the NGSC or vice versa. The first stage of the stochastic problem corresponds to the capacity booking made by the NGSC, and these capacity values remain unchanged throughout the whole process. At the next stages, the decision variables are the daily extraction amounts (and hence, the imbalances), unsatisfied demand, and the penalty cash-outs imposed by the POC.

Note that, while the POC may appear as the party with more influence in the system, the NGSC is the leader of the bilevel problem. The only reason why the NGSC is the upper level (leader) is because of the timing of the decision process. Indeed, it would seem logical that the POC, enjoying stronger control over its own facilities, has to abide to the decisions (regarding final day imbalances) that the NGSC has already made. This is because of the relative freedom that has been awarded (in the current business' practice) to the NGSC in creating imbalances to maintain healthy business in favor of its customers.

*4.2. Stochastic Model.* In [62], the authors present a bilevel multistage stochastic optimization model, which is developed to deal with a certain subsystem of the natural gas supply chain. While former models were focused on the arbitrage policies in a deterministic setting, here we have expanded the problem to include such elements as gas sales and booking costs and added a stochastic framework to model the uncertainty in demand and prices faced by the upper level decision maker (the leader).

The developed model was implemented numerically and compared to the perfect information solution (PIS) and the expected value solutions (EVS). Experimental findings show that 19 of the 21 instances deliver implementation values of over half of the PIS, whereas only one of the EVS implementation values has a relative error below 0.75. The stochastic solution implementation values are better than those of the EVS values in all but one case—which corresponds to the simplest instance tested—which testifies in favour of our approach. The performed linear reformulation also proved advantageous, as solving the original model with nonlinear levels takes considerably longer time and does not provide better solutions after up to 10 hours of running time in 20 of the 21 experiments.

Future work includes assessing the convenience of using heuristic approaches for solving the lower level (as opposed to using a specialized linear solver) and reformulating the linear lower level in the form of its duality conditions, adding these to the upper level to solve a single-level problem instead of a bilevel one. It is also worthwhile to study these models under different time series not showing seasonality is also planned, as it is the implementation of a rolling horizon approach to remedy the lack of accuracy over long-period problems (such as problem B011 involving 28 periods).

*4.3. Penalty Function Method.* Paper [64] studies a special bilevel programming problem that arises from the dealings of a Natural Gas Shipping Company (NGSC) and the Pipeline Operator Company (POC), with facilities of the latter used by the former. Because of the business relationships between these two actors, the timing and objectives of their decision-making process are different and sometimes even opposed.

In order to model that, bilevel programming was traditionally used in the above-mentioned works. Later, the problem was expanded and theoretically studied to facilitate its solution; this included extension of the upper level objective function, linear reformulation, heuristic approaches, and branch-and-bound techniques.

In this paper, the authors presented a linear programming reformulation of the latest version of the model, which is significantly faster to solve when implemented computationally. More importantly, this new formulation makes it easier to theoretically analyze the problem, allowing one to draw some conclusions about the nature of the solution of the modified problem.

When a NGSC and a POC engage in a contract, the resulting dynamics may be subject to multilevel programming analysis. In this work, an inexact penalization approach (IPA) was developed to solve the related bilevel linear programming problem, in which the NGSC is the upper level decision maker, and tries to maximize its earnings. In the meantime, the POC is the lower level decision maker trying to minimize the cash-out between both parties, while balancing the pipeline network to guarantee an adequate operation of the latter.

The IPA algorithm is adapted to the linearized versions of the problems found in [65], and theoretical work is then made to demonstrate the convergence of this solution method.

Combining the inexact penalization approach and a modified Nelder-Mead simplex algorithm has resulted in a fast and efficient enough optimization scheme, in which new iterations are generated, corrected, and then evaluated for optimality. To summarize the numerical experiments, the IPMNM approach works considerably better than both direct implementations and IPA versions without linearization. This makes a support for our linearization attempts, as well as for the advantageous usage of the IPA algorithms developed in [47]. Altogether, numerical results concerning the running time, convergence, and optimal values are presented and compared to previous reports, showing a significant improvement in speed without actual sacrifice of the solution's quality.

In conclusion, it is possible to believe that the new solution speed achieved allows one to reach a quick and more frequent balancing. Indeed, the more accurate the solution is, the more dynamic and successful the industry's response to market necessities will be.

## 5. Reduction of Upper Level Dimension in Bilevel Programming Problem

As we have already seen from the previous sections, bilevel programming modeling is a new and dynamically developing area of mathematical programming and game theory. For instance, when we study value chains, the general rule usually is that decisions are made by different parties along the chain, and these parties have often different, even opposed goals. This raises the difficulty of supply chain analysis, because regular optimization techniques (e.g., like linear programming) cannot be readily applied, so that tweaks and reformulations are often needed (*cf.* [59]).

The latter is exactly the case with the *Natural Gas Value Chain*. From extraction at the wellheads to the final consumption points (households, power plants, etc.), natural gas goes through several processes and changes ownership many a time.

Bilevel programming is especially relevant in the case of the interaction between a Natural Gas Shipping Company (NGSC) and a Pipeline Operating Company (POC). The first one owns the gas since the moment it becomes a consumption-grade fuel (usually at wellhead/refinement complexes, from now onward called the extraction points) and sells it to Local Distributing Companies (LCD), who own small, city-size pipelines that serve final costumers. Typically, NGSCs neither engage in business with end-users, nor actually handle the natural gas physically.

Whenever the volumes extracted by the NGSCs differ from those stipulated in the contracts, we say an imbalance occurs. Since imbalances are inevitable and necessary in a healthy industry, the POC is allowed to apply control mechanisms in order to avoid and discourage abusive practices (the so-called arbitrage) on part of the NGSCs. One of such tools is cash-out penalization techniques after a given operative period. Namely, if a NGSC has created imbalances in one or more pool zones, then the POC may proceed to "move" gas from positive-imbalanced zones to negative-imbalanced ones, up to the point where every pool zone has the imbalance of the same sign, that is, either all nonnegative or all nonpositive, thus rebalancing the network. At this point, the POC will either charge the NGSC a higher (than the spot) price for each volume unit of natural gas withdrawn in excess from its facilities, or pay back a lower (than the sale) price, if the gas was not extracted.

Prices as a relevant factor induce us into the area of stochastic programming instead of the deterministic approach. The formulated bilevel problem is reduced to the also bilevel one but with linear constraints at both levels (*cf.* [62]). However, this reduction involves introduction of many artificial variables, on the one hand, and generation of a lot of scenarios to apply the essentially stochastic tools, on the other hand. The latter makes the dimension of the upper level problem simply unbearable burden even for the most modern and powerful PC systems. First attempts to diminish the number of decision variables were made by the authors in [66, 67].

The aim of chapters [68, 69] is a mathematical formalization of the task of reduction of the upper level problem's dimension without affecting (if possible!) the optimal solution of the original nonlinear bilevel programming problem. Under a couple of quite reasonable assumptions about the data of the original bilevel programming problem, the authors of [68, 69] established that the modified problem obtained by translating part of upper level variables to the lower level and replacing the original lower level program with an appropriate equilibrium problem will have the same solution set as the original bilevel program.

A bit more specialized and profound results were deduced in [68] for the linear bilevel program by making use of certain tools from the previous works [70–74]. As paradoxically it could sound, in the linear case, the problem is much more complicated. Indeed, the uniqueness of a generalized Nash equilibrium (GNE) at the lower level of is much too restrictive a demand. As was shown by Rosen [72], the uniqueness of a so-called normalized GNE is rather more realistic assumption. This idea was further developed later by many authors, including the authors of [69, 73].

Following the line proposed in [72], the authors of [69] introduce and study the concept of *normalized generalized Nash equilibrium* (NGNE) defined similarly to the concept from [72]. Based upon the revealed properties of such a entity, they establish the existence and uniqueness results for the lower level problem. Hence, the coincidence of the solution sets of the original bilevel (linear or nonlinear) program and the modified model obtained by the translation of part of variables from the upper to the lower level is demonstrated.

To conlcude, chapters [68, 69] deal with an interesting problem of reducing the number of variables at the upper level of bilevel programming problems. The latter problems are widely used to model various applications, in particular, the natural gas cash-out problems described in [59, 62]. To solve these problems with stochastic programming tools, it is important that part of the upper level variables be governed at the lower level, to reduce the number of (upper level) variables, which are involved in generating the scenario trees.

The chapters present certain preliminary results recently obtained in this direction. In particular, it has been demonstrated that the desired reduction is possible when the lower level optimal response is determined uniquely for each vector of upper level variables. In [69], the necessary base for similar results is prepared for the general case of bilevel programs with linear constraints, when the uniqueness of the lower level optimal response is quite a rare case. However, if the optimal response is defined for a fixed set of Lagrange multipliers, then it is possible to demonstrate (following the ideas and techniques from [72]) that the so-called *normalized* Nash equilibrium is unique. The latter gives one a hope to get the positive results about reducing the dimension of the upper level problem *without* affecting the solution of the original bilevel programming problem.

## 6. Allocation Models as Bilevel Programming Problems

Bilevel programming has also served as a suitable option for modeling allocation problems where two-hierarchized levels with different objectives are involved. At each level, the decision maker aims to optimize his own interest. The predefined existing hierarchy allows that the upper level has complete information about the lower level's decision on the allocation, but not on the vice versa manner. In particular, bilevel programming offers a convenient framework for dealing with the allocation problems.

An important and very common problem that appears in these kinds of situations is the allocation of resources or the allocation of parties in the whole process considered. Hence, we are going to divide this literature review in two directions: first, the previous works done where the optimal allocation of resources are described, and then, the papers related to optimally allocate customers, distribution centers, plants, or other parties involved in a specific supply chain are refereed.

*6.1. Bilevel Allocation of Resources.* When considering a company's personnel and workers as *limited resources*, we could mention paper [75] where the main department boasting several branching divisions needs to allocate the personnel (workers, technicians, and management personnel) for the company's tasks. The leader intends to maximize its benefit by allocating the specific workers to the divisions, while the followers aim to maximize their own benefits using the assigned personnel. The authors of [75] solved the proposed model by applying a *simulation bionic* algorithm. The main issue is that they did not make any conclusions about the quality of the thus obtained solution due to the complexity inherent to the bilevel model.

In [76], the minimum total time for finishing jobs in a system is sought. In that problem, the leader is the job scheduler who tries to optimize the system performance by allocating the workers to the machines. On the other hand, the follower is represented by many noncooperative workers seeking to use a set of common machines minimizing the latency of their work schedule. Three polynomial-time algorithms for solving the problem are proposed, and complexity results are given demonstrating that this problem is NP-hard.

Next, one can find a plenty of papers devoted to the analyzis of the allocation of water to different regions of the world. For example, in [77] a nonlinear bilevel programming model with fuzzy random variables for distributing (in an equitable way) the water in a region is studied. The whole community (society) is seen as the leader, and the followers are seen as the subareas contained in the region. Both decision levels strive to maximize their economic gain. The authors of [77] proposed a hybrid heuristic based on an interactive fuzzy programming technique and a genetic algorithm. Also, an application to a real case study was made showing the reasonable performance of the developed solution method.

Paper [78] examines a similar situation: a bilevel multiobjective linear programming model is considered. It is important to note that the lower level problem contains multiple objective functions. The leader has to allocate the amount of water destined for irrigation, industry, domesticity, and ecology in order to maximize the benefit for the region. Then, the follower optimizes its gain using the water resource doomed for each purpose. The problem is solved by using fuzzy goal programming in the upper level and a tolerance-based approach in the lower level. Their model and methodology was validated in a case study from China. In [79], more references concerning this particular topic can be found (in Japanese).

Another interesting application is about housing allocation. In [80], this problem for a continuum transportation system is analyzed. The leader selects the optimal housing development pattern while the follower decides about the allocation of the houses based on their renting and travel costs. The lower level problem is defined by a set of differential equations and it is solved by the finite element method. The results obtained from numerical experimentation show that the algorithm seems to be efficient enough. An extension of the previous work was done in [81]. The main difference is in that the leader optimizes the housing allocation in order to achieve the minimum $CO_2$ emissions, while the followers aim is to find the equilibrium among the users in a transportation system. The authors of [81] also adapted the finite element method and proposed two alternative solution algorithms based on the Newton-Raphson procedure and the convex combination approach. The computational tests showed that traffic intensity, $CO_2$ emissions, and transport demand are balanced along with the best housing allocation.

Bilevel programs related to the optimal allocation of a specific product can also be found in the literature. For example, [82] presents a problem where a company markets products and allocate resources to two producer factories that consume the resources. Hence, the model can be viewed and treated as a Stackelberg equilibrium problem, because in the lower level, both followers compete for the common allocated resources trying to optimize their own criteria. A hybrid intelligent algorithm based on fuzzy simulation, as well as neural network, and genetic algorithms are proposed for solving this bilevel problem.

Wang and Lootsma [83] introduced a bilevel model for the case when the general manager tries to allocate resources among the different departments of the company. In the upper level, the correct allocation of the resources to the departments is made in order to maximize the company's total revenue. On the other hand, in the lower level, each department estimates its own benefit generated with the allocated resources. A numerical example is given to illustrate the proposed exact method.

As we mentioned before, bilevel programming allows a realistic mathematical modeling for a very wide application areas. We are going to confirm this fact with the work done by Burgard et al. [84] where a genomic problem is addressed. In that problem, the leader maximizes the bioengineering objective, that is, the chemical production, and the follower optimizes the flux allocation based on the biomass generated through the gene deletions.

*6.2. Bilevel Allocation for the Supply Chain Models.* It is well known that supply chains involve many components in the whole process. At some point of the supply chain, an allocation is required, for example, to allocate customers to plants, demanded zones to distribution centers, vehicles to producers, and so forth.Under this scheme, Calvete et al. [85] introduced a production-distribution bilevel problem, in which a company (the leader) is dedicated exclusively to the allocation of customers to distribution centers satisfying their demand of products. Another company (the follower) is doomed to produce these products. The leader will distribute the products and purchase them from some plants, and then the distribution centers will transport them to their customers meeting their requirements in order to minimize the distribution costs. On the other hand, the follower decides its own production plan based on the production capacity of the plants and by considering the requirements of the demand grouped in the distribution centers seeking to minimize the operation costs. The authors of [85] considered a real case from a company in Spain and also some benchmark instances. Furthermore, they solved this problem by using a heuristic algorithm based on an ant colony optimization method delivering pretty good quality solutions in a reasonable time.

Also, Legillon et al. [86] considered the same problem proposing a coevolutionary algorithm without improving the solution quality given in the seminal paper. Camacho-Vallejo et al. [87] developed a method based on scatter search obtaining the best known results for the benchmark instances. In [88], a single-commodity, multiplant network with multiple depots is studied. The leader seeks to minimize the total cost (i.e., the cost associated with the distribution from the plants to depots and then to the customers, plus the warehousing costs and the operation costs of the depots) of locating depots and allocating customers to them. The follower intends to balance the workload of the system improving the customer service and finding a trade-off between cost and efficiency. A standard genetic algorithm was proposed [88] in order to solve some randomly generated test problems with demonstrating certain opportunity areas for improving its performance.

Humanitarian logistics have given rise to application of bilevel programming frameworks for dealing with situations that appear in that area. Feng and Wen [89] considered the bilevel program where an earthquake affected the local transportation network. Here, the leader tries to maximize the flow of vehicles entering the affected area to provide assistance, whereas the followers attempt to travel through an unaffected route to minimize their total travel time. Since this situation generates traffic jams and negatively impacts the recovery and relief efforts, a government agency regulates the use of existing roads. In order to solve the proposed model, a genetic algorithm was implemented and validated in a case study showing that this algorithm is an effective tool to solve the problem in question.

In their turn, Wang et al. [90] proposed a model for locating storage centers and allocating the sent aid. The leader minimizes the cost of locating the storage centers, allocation of sent aid, distribution, and penalties, while the follower (an affected community) optimizes its own cost based on the resources allocated to each community. A small test instance was created for testing the developed particle swarm optimization algorithm showing the ease of its implementation.

Similar to the models discussed above, Sun et al. [91]. seek an optimal decision about locating distribution centers by the search of an equilibrium among the customers' costs. The leader will locate new distribution centers to minimize fixed and variable costs while meeting the demand by a set of customers. In its turn, the follower will allocate the customers to the distribution centers so as to minimize the cost of meeting their demand. An algorithm that exploits the special structure of the lower level problem and a *branch-and-bound* (B&B) scheme in the upper level is proposed to deal with this bilevel program. In a different context (but with a similar structure) Xu and Wei [92] modeled a problem related to the waste management of constructions and demolitions. The government is the leader that has to make the decision about locating the waste collection depots and processing centers. The administrators of different construction waste management systems control the allocation of the waste to the located facilities. Both objectives functions minimize their own costs in a fuzzy random environment. An improved particle swarm optimization algorithm was designed to treat and solve the latter problem.

It seems that facility location and customers' allocation requirements can be effectively modeled with bilevel programming when taking into account the customers' demand at the facilities that will serve them. Various papers in which the customers are allocated to the facilities according to a predetermined list of preferences can be found in the literature; *see* [93–96]. In all those papers, the facility location problem under customers' preferences is studied. In the bilevel program induced, the leader has to locate some facilities, while the follower will allocate the customers optimizing their preestablished preferences towards the facilities. The first three papers (i.e., [93–95]) developed valid two-sided bounds for the objective functions invloved in this problem, and the last two works (i.e., [95, 96]) implemented heuristic algorithms to process the bilevel model.

Moreover, competitive facility location models have been approached with bilevel programming, too. In that problem, two competing firms have to locate some facilities in order to capture the maximum demand of the existing customers. With an aim to classify the problem as a bilevel program, a hierarchy among the firms must exist in the model. A lot of variations of these models have been published. The differentiation relies on the customers' behavior; for example, the customers may be allocated to the facilities based on a predefined criterion, such as the shortest distance, a list of preferences, preestablished contracts, or in a random way. Another important factor is the characteristics of the competing firms, for instance, (i) if they have an exact number of facilities to be located, that is, $(r \mid p)$-centroid problem; (ii) whether one firm already has located facilities and the other firm has to locate new ones, that is, $(r, X_p)$-medianoid problem. The existence of facilities, the possibility of closing some or make them more attractive, and so forth,—all them are the issues that are addressed in these models. It is important to note that in competitive facility location problems, neither the leader nor the follower will make the decision of the customers' allocation, but this allocation implicitly appears in the process and clearly affects both levels' decisions. The reader is referred to [97–106] in order to have a closer look to particular models in this area.

Further, the design of telecommunication networks has also been analyzed as a bilevel programming scheme. A problem within this area is the one studied by Kim et al. [107], in which the topological design of a local area network is proposed. The problem consists of allocating users to clusters and the union of clusters by bridges in order to obtain a minimum response time network boasting at the same time the minimum connection costs. Therefore, the decision concerning the optimal allocation of users to clusters will be made by the leader, while the follower will make the decision about connecting all the clusters by forming a spanning tree. The authors [107] applied a coevolutionary genetic algorithm based on Nash equilibrium to solve the problem.

Finally, optimization in ports has also attracted the attention of researches and found applications of bilevel programming: *compare* Lee et al. [108], where a problem for scheduling berth and quay cranes is studied. In that problem, the leader deals with the berth allocation problem minimizing the sum of waiting and handling times of each vessel. On the other hand, the follower solves the quay crane scheduling problem in order to minimize the total time until all the vessels and the quay cranes have finished up their activities. Owing to the difficulty of the exact solution of this bilevel problem, a genetic algorithm that finds reasonable quality solutions is proposed in [108].

## 7. Information Protection and Cybersecurity Problems as Bilevel Programs

The methods and approaches solving bilevel programming problems also are actual in the areas of information protection and cybersecurity. However, the solutions in these similar areas have some special features, especially, concerning certain cryptographic applications. One of these features is the following. There exists a standard (single-level) mathematical formalization of the cryptographic problem, but it has been shown to bear some flaws. Thus, the proposed methods and approaches based on the bilevel programming techniques helps eliminate those deficiences and enhance the processing of problems of the information protection and cybersecurity on the new quality level.

At the same time, the problems of information protection and cybersecurity clealy lack good interpretations with the help of the bilevel programming apparatus. Therefore, this section of the survey is presenting the first-time and original review of the possible treatment of these important information protection and cybersecurity problems as *bilevel programs*.

*7.1. Some Cryptographic Applications.* One of the urgent problems of public key cryptosystem improvements is the increase of the qulity of software performance and hardware implementations. One of the approaches helping improve the functioning of cryptosystems is marking up the performance of finite field arithmetic concerning operations of multiplication. A possible way to do is to widely apply the bilevel programming techniques.

As the well-known publications show (*cf.* [109–112], to mention only few), the most effective multiplication algorithms have been provided by Comba [109] and Karatsuba [112]. However, Comba's algorithm shows somewhat better results in numerous rendition (benchmark) tests of software implementations on modern platforms. The combined Karatsuba-Comba multiplication (KCM) algorithm for processors of the reduced instruction set computers (RISC-processors) is described in paper [113].

The KCM-algorithm is an example of a promising combination of those by Comba and Karatsuba, while Karatsuba's algorithm is especially often used for the machine word multiplication. As a result, the main goal of that paper [113] is to provide a suggestion for the effective increasing of software implementation of the finite field $GF(p)$ multiplication (squaring) with the aid of Comba's algorithm. Such research was motivated by the necessity to obtain the effective confirmation of software implementation of some known algorithms for continuous development of the modern 32-bit and 64-bit platforms. It is important to mention that the last ten years have seen a rapid development of multicore processors and multiprocessor systems [113].

*7.2. Software Implementation.* With the recent boost of information technology in modern society, the problem of information security became of special urgency. The most difficult task is to provide secure handling and storage of critical and confidential data for government and private companies, banks, and other systems. A solution to this problem is to implement systems that provide for information confidentiality, integrity, authenticity, and accessibility by means of cryptographic software and cryptographic hardware based on some approaches making use of bilevel programming.

At the same time, cryptoanalytical methods taking advantage of the progress in capabilities of modern computers demand high requirements on the security parameters of modern cryptosystems with the use of the well-known techniques and devices of bilevel programming. Moreover, the increased amount of data processed in modern information systems needs a quite high-level performance of the modern cryptosystems. Hence, the timing requirements to cryptographic applications have increased dramatically; that is, prospective cryptoalgorithms must provide efficient processing of bulk data when applying bilevel programming and, at the same time, a high level of security.

So far, most research activity has been carried out about some theoretical aspects of hyperelliptic curve cryptosystems (HECC), including many improvements of the underlying arithmetic of the hyperelliptic curves. On the implementation side, improvements for specific processors and hardware platforms have been analyzed. The present approach provides a very important contribution towards practical implementation of HECC by showing how to build an efficient hyperelliptic curve of digital signature algorithm (HECDSA) implementation and provides cryptographically suitable curves. Unfortunately, the published results on practical implementations of HECC are rare (*see*, for example, [114, 115]). This solution is intended to provide very practical facts for the implementation of an HECDSA system with all its necessary details at the interpretation with the help of the bilevel programming techniques. There are numerous modern publications dealing with HECC, but they describe no validated system parameters for the efficient implementation of a workable cryptosystem.

The lack of publications dedicated to this topic was the motivation behind the thorough summary of all results for efficient HECC implementation presented in this review and the comparison of HECC (HECDSA) with the existing elliptic curve cryptosystems (ECC) and/or elliptic curve of digital signature algorithms (ECDSA) based on the use of some bilevel programming methods.

*7.3. Cybersecurity Applications.* The bilevel formulation is investigated through a problem in which the goal of the destructive agent is to minimize the number of power system components that must be destroyed in order to cause a loss of load greater than or equal to a specied level. This goal is tempered by the logical assumption that, following a deliberate outage, the system operator will implement all feasible corrective actions to minimize the level of system load shed.

The resulting nonlinear mixed-integer bilevel programming formulation is transformed into an equivalent single-level mixed-integer linear program by replacing the lower level optimization problem with its Karush-Kuhn-Tucker (KKT) optimality conditions and also converting a number of nonlinearities to linear equivalents using some well-known integer algebra results. The equivalent formulation has been tested in [116] on two case studies, including the 24-bus IEEE reliability test system (RTS) through the use of commercially available software.

The bilevel model specically allows one to dene different objective functions for the terrorist and the system operator and permits to impose constraints on the upper level optimization problem. The latter are functions of both the upper and lower level variables. This degree of exibility is not possible to implement through the existing *max-min* models.

As present, researchers have begun to look into some new ways of addressing the security assessment problem, here called the *Terrorist Threat Problem* (TTP). For example, in [117], a multiagent system was proposed capable of assessing power system vulnerability, monitoring hidden failures of protection devices, and providing adaptive control actions to prevent catastrophic failures and cascading sequences of events.

*Attack tree* (AT) is another widely used combinatorial model in the cybersecurity analysis. The basic formalism of AT does not take into account defense mechanisms. *Defense trees* (DT) have been developed to investigate the effect of defense mechanisms using measures such as attacker's cost and security cost, return on investment (ROI) and return on attack (ROA). DT, however, places defense mechanisms only at the leaf node level while the corresponding ROI/ROA analysis does not incorporate the probability of attack. In an *attack response tree* (ART), an attacker-defender game was used to find an optimal policy from the countermeasures' pool. The latter suffers from the problem of state-space explosion, since a solution in ART is sought by means of a *partially observable stochastic game* model. In [118], the authors have presented a novel attack tree named the *attack countermeasure tree* (ACT), in which (i) defense mechanisms can be applied at any node of the tree, not just at the leaf node level; (ii) some qualitative analysis (using *min-cuts*, structural and Birnbaum importance measures) and probabilistic analysis (using attacker's and security costs, the system risk,the impact of an attack, ROI, and ROA) can be performed; (iii) the optimal countermeasure set can be selected from the pool of defense mechanisms without constructing a state-space model. They have used single- and multiobjective optimization tools to find suitable countermeasures under different constraints. In addition, they have illustrated the features of ACT using a practical case study, namely, a supervisory control and data acquisition (SCADA) attack.

Finally, some authors [119] have proposed a *trilevel* model. Cybersecurity is becoming an area of growing concern in the electric power industry with the development of *smart grid*. A false data injection attack, which is against the state estimation through a SCADA network, has recently attracted the ever wider interest of researchers. This review [119] further develops the concept of a *Load redistribution* (LR) attack, a special type of the false data injection attack. The damage from LR attacks to power system operations can manifest in an immediate or a delayed fashion. For the immediate attacking goal, they have shown in [119] that the most damaging attack can be identified through a *max-min* attacker-defender model. Benders decomposition within a restart framework is used to solve the bilevel immediate LR attack problem with a moderate computational effort. Its efficiency has been validated by the Karush-Kuhn-Tucker-(KKT-) based method solution in their previous work. For the delayed attacking goal, the authors of [119] have proposed a *trilevel* model to identify the most damaging attack and transform the model into an equivalent single-level mixed-integer problem for its final solution. In order to summarize, the techniques developed in [119] enable a quantitative analysis of the damage from LR attacks to the power system operations and security and hence provide an in-depth insight into an effective attack prevention when resources (budgets) are limited. A 14-bus system is used to test the correctness of the proposed model and algorithm.

## 8. Concluding Remarks

In this paper, we present a survey of Bilevel Programming and Application area, closely related to applied problems such as natural gas imbalance cash-out problem, toll optimization problem, and others. Recent results and trends in the mixed-integer bilevel programming models with linear objective function and constraints are also described.

Many open questions still exist in Bilevel Programming theory, especially in relation to applications. New topics/questions arise as, for example, application of non-smooth/variational analysis. Many new applications are found; much is yet open with respect to solution algorithms; important are also mixed-discrete bilevel optimization problems. All these items have not been included in this survey only due to the space limitations, but we hope to enlight them in the nearest future.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The research activity of the second author was financially supported by the R&D Department (Cátedra de Investigación) CAT-174 of the Tecnológico de Monterrey (ITESM), Campus Monterrey and by the SEP-CONACYT Projects CB-2008-01-106664 and CB-2013-01-221676, Mexico. Also, the work of the fourth author was supported by the National Council of Science and Technology (CONACyT) of Mexico as part of the Projects CB-2011-01-169765, PROMEP 103.5/11/4330, and PAICYT 464-10. The research activity of the fifth author was financially supported by the Project PROMEP/103.5/10/3889, the Academic Groups Research Project PROMEP/103.5/12/4953, and the Autonomous University of Nuevo León (UANL) within the Support Program for Scientific Research and Technology (PAICYT) with the Project CE960-11. The coauthors would also like to express their profound gratitude to the two anonymous referees whose comments and suggestions have helped improve the paper essentially.

## References


[1] S. Dempe, *Foundations of Bilevel Programming*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[2] S. Dempe, B. S. Mordukhovich, and A. B. Zemkoho, "Necessary optimality conditions in pessimistic bilevel programming," *Optimization*, vol. 63, no. 4, pp. 505–533, 2014.

[3] S. Dempe and J. Dutta, "Is bilevel programming a special case of a mathematical program with complementarity constraints?" *Mathematical Programming A*, vol. 131, no. 1-2, pp. 37–48, 2012.

[4] S. Dewez, M. Labb, P. Marcotte, and G. Savard, "New formulations and valid inequalities for a bilevel pricing problem," *Operations Research Letters*, vol. 36, no. 2, pp. 141–149, 2008.

[5] L. Thi, T. Duc, and P. Dinh, "A DC programming approach for a class of bilevel programming problems and its application in portfolio selection," *Numerical Algebra, Control and Optimization*, vol. 2, no. 1, pp. 167–185, 2012.

[6] L. N. Vicente and P. H. Calamai, "Bilevel and multilevel programming: a bibliography review," *Journal of Global Optimization*, vol. 5, no. 3, pp. 291–306, 1994.

[7] W. Wiesemann, A. Tsoukalas, P.-M. Kleniati, and B. Rustem, "Pessimistic bi-level optimisation," Tech. Rep., Imperial College London and Massachusetts Institute of Technology, 2012.

[8] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer, *Non-Linear Parametric Optimization*, Birkhäauser, Basel, Switzerland, 1983.

[9] R. Lucchetti, F. Mignanego, and G. Pieri, "Existence theorems of equilibrium points in Stackelberg games with constraints," *Optimization*, vol. 18, no. 6, pp. 857–866, 1987.

[10] D. Fanghänel, *Zwei-Ebenen-Optimierung mit Diskreter Unterer Ebene und Stetiger Oberer Ebene*, TU Bergakademie Freiberg, Freiberg, Germany, 2006.

[11] H. von Stackelberg, *Marktform und Gleichgewicht*, Julius Springer, Vienna, Austria, 1934.

[12] J. Bracken and J. T. McGill, "Mathematical programs with optimization problems in the constraints," *Operations Research*, vol. 21, pp. 37–44, 1973.

[13] W. Candler and R. Norton, "Multilevel programming and development policy," Tech. Rep. 258, World Bank Staff, Washington, Wash, USA, 1977.

[14] L. N. Vicente, C. A. Floudas, and P. M. Pardalos, "Bilevel programming: introduction, history and overview," in *Encyclopedia of Optimization*, C. A. Floudas and P. M. Pardalos, Eds., Kluwer Academic, Dordrecht, The Netherlands, 2001.

[15] J. F. Bard, *Practical Bilevel Optimization: Algorithms and Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[16] K. Shimizu, Y. Ishizuka, and J. F. Bard, *Nondifferentiable and two-level mathematical programming*, Kluwer Academic Publishers, Boston, Mass, USA, 1997.

[17] S. Dempe and V. V. Kalashnikov, Eds., *Optimization with Multivalued Mappings: Theory, Applications and Algorithms*, Springer, Berlin, Germany, 2006.

[18] A. Migdalas, P. M. Pardalos, and P. Värbrand, Eds., *Multilevel Optimization: Algorithms and Applications*, Kluwer Academic, Dordrecht, The Netherlands, 1998.

[19] S. Dempe, "Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints," *Optimization*, vol. 52, no. 3, pp. 333–359, 2003.

[20] W. Candler and R. Townsley, "A linear two-level programming problem," *Computers & Operations Research*, vol. 9, no. 1, pp. 59–76, 1982.

[21] Z. Luo, J. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK, 1996.

[22] J. Outrata, M. Kočvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[23] V. Demiguel, M. P. Friedlander, F. J. Nogales, and S. Scholtes, "A two-sided relaxation scheme for mathematical programs with equilibrium constraints," *SIAM Journal on Optimization*, vol. 16, no. 2, pp. 587–609, 2005.

[24] S. Leyffer, G. López-Calva, and J. Nocedal, "Interior methods for mathematical programs with complementarity constraints," *SIAM Journal on Optimization*, vol. 17, no. 1, pp. 52–77, 2006.

[25] M. Kojima, "Strongly stable stationary solutions in nonlinear programs," in *Analysis and Computation of Fixed Points*, S. M. Robinson, Ed., vol. 43, pp. 93–138, Academic Press, New York, NY, USA, 1980.

[26] D. Ralph and S. Dempe, "Directional derivatives of the solution of a parametric nonlinear program," *Mathematical Programming*, vol. 70, no. 2, pp. 159–172, 1995.

[27] S. Scholtes, "Introduction to piecewise differentiable equations," Tech. Rep. 53/1994, Universität Karlsruhe, Institut für Statistik und Mathematische Wirtschaftstheorie, 1994, http://www.eng.cam.ac.uk/~ss248/publications/index.html.

[28] S. Dempe, "A necessary and a sufficient optimality condition for bilevel programming problems," *Optimization*, vol. 25, no. 4, pp. 341–354, 1992.

[29] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, Vol. 1: Basic Theory*, Springer, Berlin, Germany, 2006.

[30] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation*, vol. 2, Springer, Berlin, Germany, 2006.

[31] R. T. Rockafellar and R. J. Wets, *Variational Analysis*, Springer, Berlin, Germany, 1998.

[32] J. V. Outrata, "Necessary optimality conditions for Stackelberg problems," *Journal of Optimization Theory and Applications*, vol. 76, no. 2, pp. 305–320, 1993.

[33] J. J. Ye and D. Zhu, "New necessary optimality conditions for bilevel programs by combining the MPEC and value function approaches," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1885–1905, 2010.

[34] S. Dempe, J. Dutta, and B. S. Mordukhovich, "New necessary optimality conditions in optimistic bilevel programming," *Optimization*, vol. 56, no. 5-6, pp. 577–604, 2007.

[35] Y. Ishizuka and E. Aiyoshi, "Double penalty method for bilevel optimization problems," *Annals of Operations Research*, vol. 34, no. 1–4, pp. 73–88, 1992.

[36] A. G. Mersha and S. Dempe, "Linear bilevel programming with upper level constraints depending on the lower level solution," *Applied Mathematics and Computation*, vol. 180, no. 1, pp. 247–254, 2006.

[37] F. Nožička, J. Guddat, H. Hollatz, and B. Bank, *Theorie der Linearen Parametrischen Optimierung*, Akademie, Berlin, Germany, 1974.

[38] Energy Information Administration, *FERC Order 636: The Restructuring Rule*, 2005, http://www.eia.gov/forecasts/steo/.

[39] Energy Information Administration, "FERC Policy on System Ownership Since 1992," http://www.eia.doe.gov/oil_gas/natural_gas/analysis_publications/ngmajorleg/fercpolicy.html.

[40] IHS Engineering, "EC Proposes New Legislat ion for European Energy Policy," 2008, http://engineers.ihs.com/news/eu-en-energy-policy-9-07.html.

[41] Environmental Protection Agency, "The Impacts of FERC Order 636 on coal mine gas project development," 2008, http://www.epa.gov/cmop/docs/pol004.pdf.

[42] K. T. Midthun, *Optimization Models for Liberalized Natural Gas Markets*, Norwegian University of Science and Technology (NTNU), Faculty of Social Science and Technology Management, Department of Industrial Economics and Technology Management, Trondheim, Norway, 2008.

[43] S. Dempe, V. V. Kalashnikov, and R. Z. Ríos-Mercado, "Discrete bilevel programming: application to a natural gas cash-out problem," *European Journal of Operational Research*, vol. 166, no. 2, pp. 469–488, 2005.

[44] M. J. Doane and D. F. Spulber, "Open access and the evolution of the US spot market for natural gas," *Journal of Law and Economics*, vol. 34, no. 2, pp. 447–517, 1994.

[45] R. Gutiérrez, A. Nafidi, and R. G. Sánchez, "Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model," *Applied Energy*, vol. 80, no. 2, pp. 115–124, 2005.

[46] F. K. Lyness, "Gas demand forecasting," *Journal of the Royal Statistical Society D*, vol. 33, no. 1, pp. 9–12, 1984.

[47] V. V. Kalashnikov and R. Z. Ríos-Mercado, "A natural gas cash-out problem: a bilevel programming framework and a penalty function method," *Optimization and Engineering*, vol. 7, no. 4, pp. 403–420, 2006.

[48] N. Keyaerts, L. Meeus, and W. D'Haeseleer, "Analysis of balancing-system design and contracting behavior in the natural gas markets," in *European Doctoral Seminar on Natural Gas Research*, vol. 24, Delft, The Netherlands, 2009.

[49] H. G. Huntington, "Federal price regulation and the supply of natural gas in a segmented field market," *Land Economics*, vol. 54, no. 3, pp. 337–347, 1978.

[50] A. Tomasgard, F. Romo, M. Fodstad, and K. T. Midthun, "Optimization models for the natural gas value chain," in *Geometric Modeling, Numerical Simulation, and Optimization*, Applied Mathematics at SINTEF, Chapter Optimization Models for the Natural Gas Value Chain, Springer, 2007.

[51] C. Borraz-Sánchez and R. Z. Ríos-Mercado, "A hybrid metaheuristic approach for natural gas pipeline network optimization," in *Hybrid Metaheuristics*, vol. 3636 of *Lecture Notes in Computer Science*, pp. 54–65, Springer, New York, NY, USA, 2005.

[52] A. Chebouba, F. Yalaoui, A. Smati, L. Amodeo, K. Younsi, and A. Tairi, "Optimization of natural gas pipeline transportation using ant colony optimization," *Computers and Operations Research*, vol. 36, no. 6, pp. 1916–1923, 2009.

[53] A. Kabirian and M. R. Hemmati, "A strategic planning model for natural gas transmission networks," *Energy Policy*, vol. 35, no. 11, pp. 5656–5670, 2007.

[54] S. A. Gabriel, J. Zhuang, and S. Kiet, "A large-scale linear complementarity model of the North American natural gas market," *Energy Economics*, vol. 27, no. 4, pp. 639–665, 2005.

[55] R. Egging, S. A. Gabriel, F. Holz, and J. Zhuang, "A complementarity model for the European natural gas market," *Energy Policy*, vol. 36, no. 7, pp. 2385–2414, 2008.

[56] D. Hawdon, "Efficiency, performance and regulation of the international gas industry: a bootstrap DEA approach," *Energy Policy*, vol. 31, no. 11, pp. 1167–1178, 2003.

[57] K. G. Arano and B. F. Blair, "An ex-post welfare analysis of natural gas regulation in the industrial sector," *Energy Economics*, vol. 30, no. 3, pp. 789–806, 2008.

[58] B. Esnault, "The need for regulation of gas storage: the case of France," *Energy Policy*, vol. 31, no. 2, pp. 167–174, 2003.

[59] V. V. Kalashnikov and R. Z. Ríos-Mercado, "A penalty-function approach to a mixed-integer bilevel programming problem," in *Proceedings of the 3rd International Meeting on Computer Science*, C. R. Zozaya, Ed., vol. 2, pp. 1045–1054, Aguascalientes, Mexico, September 2001.

[60] S. Dempe, V. V. Kalashnikov, and G. A. Pérez-Valdés, "Mixed-integer bilevel programming: application to an extended gas cash-out problem," in *Proceedings of the International Business and Economics Research Conference (IBERC & TLC '06)*, p. 14, Las Vegas, Nev, USA, 2006.

[61] V. V. Kalashnikov, N. I. Kalashnykova, and G. A. Pérez, "Natural gas cash-out problem with price predictions," in *Proceedings of the International Applied Business Research Conference (ABRC '07)*, R. D. Clute, Ed., p. 13, Mazatlán, Mexico, 2007.

[62] V. V. Kalashnikov, G. A. Pérez-Valdés, A. Tomasgard, and N. I. Kalashnykova, "Natural gas cash-out problem: bilevel stochastic optimization approach," *European Journal of Operational Research*, vol. 206, no. 1, pp. 18–33, 2010.

[63] P. Kall and S. W. Wallace, *Stochastic Programming*, John Wiley & Sons, Chichester, UK, 1994.

[64] S. Dempe, V. V. Kalashnikov, G. A. Pérez-Valdés, and N. Kalashnykova, "Natural gas bilevel cash-out problem: convergence of a penalty function method," *European Journal of Operational Research*, vol. 215, no. 3, pp. 532–538, 2011.

[65] V. V. Kalashnikov, G. A. Pérez, and N. I. Kalashnykova, "A linearization approach to solve the natural gas cash-out bilevel problem," *Annals of Operations Research*, vol. 181, no. 1, pp. 423–442, 2010.

[66] V. V. Kalashnikov, T. I. Matis, and G. A. Pérez-Valdés, "Time series analysis applied to construct US natural gas price functions for groups of states," *Energy Economics*, vol. 32, no. 4, pp. 887–900, 2010.

[67] V. V. Kalashnikov, G. A. Pérez-Valdés, T. I. Matis, and N. I. Kalashnykova, "US natural gas market classification using pooled regression," *Mathematical Problems in Engineering*, vol. 2014, Article ID 695084, 9 pages, 2014.

[68] V. V. Kalashnikov, S. Dempe, G. A. Pérez-Valdés,, and N. I. Kalashnykova, "Reduction of dimension of the upper level problem in a bilevel programming model. Part 1," in *Advances in Intelligent Decision Technologies*, J. Watada, J. Phillips-Wren, G. Jain, and R. J. Howlett, Eds., vol. 10 of *Smart Innovation, Systems and Technologies*, pp. 255–264, Springer, Heidelberg, Germany, 2011.

[69] V. V. Kalashnikov, S. Dempe, G. A. Pérez-Valdés, and N. I. Kalashnykova, "Reduction of dimension of the upper level problem in a bilevel programming model. Part 2," in *Advances in Intelligent Decision Technologies*, J. Watada, G. Phillips-Wren, L. C. Jain, and R. J. Howlett, Eds., vol. 10 of *Smart Innovation, Systems and Technologies*, pp. 265–272, Springer, Heidelberg, Germany, 2011.

[70] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York, NY, USA, 1980.

[71] O. L. Mangasarian, "Uniqueness of solution in linear programming," *Linear Algebra and Its Applications*, vol. 25, pp. 151–162, 1979.

[72] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave N-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[73] R. Nishimura, S. Hayashi, and M. Fukushima, "Robust Nash equilibria in $N$-person non-cooperative games: uniqueness and reformulation," *Pacific Journal of Optimization*, vol. 5, no. 2, pp. 237–259, 2009.

[74] G. K. Saharidis and M. G. Ierapetritou, "Resolution method for mixed integer bi-level linear problems based on decomposition technique," *Journal of Global Optimization*, vol. 44, no. 1, pp. 29–51, 2009.

[75] L. Lei, C. Guang-Nian, and L. Chen-Xin, "Research on problems bilevel programming for personnel allocation in enterprise," in *Proceedings of the 17th International Conference on Management Science & Engineering (ICMSE '10)*, pp. 293–298, Melbourne, Australia, November 2010.

[76] T. Roughgarden, "Stackelberg scheduling strategies," *SIAM Journal on Computing*, vol. 33, no. 2, pp. 332–350, 2004.

[77] J. Xu, Y. Tu, and Z. Zeng, "Bilevel optimization of regional water resources allocation problem under fuzzy random environment," *Journal of Water Resources Planning and Management*, vol. 139, no. 3, pp. 246–264, 2013.

[78] S. Fang, P. Guo, M. Li, and L. Zhang, "Bilevel multiobjective programming applied to water resources allocation," *Mathematical Problems in Engineering*, vol. 2013, Article ID 837919, 9 pages, 2013.

[79] G. Hongli, L. Juntao, and G. Hong, "A survey of bilevel programming model and algorithm," in *Proceedings of the 4th International Symposium on Computational Intelligence and Design (ISCID '11)*, pp. 199–203, Hangzhou, China, October 2011.

[80] H. W. Ho and S. C. Wong, "Housing allocation problem in a continuum transportation system," *Transportmetrica*, vol. 3, no. 1, pp. 21–39, 2007.

[81] J. Yin, S. C. Wong, N. N. Sze, and H. W. Ho, "A continuum model for housing allocation and transportation emission problems in a polycentric city," *International Journal of Sustainable Transportation*, vol. 7, no. 4, pp. 275–298, 2013.

[82] R. Liang, J. Gao, and K. Iwamura, "Fuzzy random dependent-chance bilevel programming with applications," in *Advances in Neural Networks-ISNN 2007*, vol. 4492 of *Lecture Notes in Computer Science*, pp. 257–266, Springer, Berlin, Germany, 2007.

[83] S. Y. Wang and F. A. Lootsma, "A hierarchical optimization model of resource allocation," *Optimization*, vol. 28, no. 3-4, pp. 351–365, 1994.

[84] A. P. Burgard, P. Pharkya, and C. D. Maranas, "OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization," *Biotechnology and Bioengineering*, vol. 84, no. 6, pp. 647–657, 2003.

[85] H. I. Calvete, C. Galé, and M. Oliveros, "Bilevel model for productiondistribution planning solved by using ant colony optimization," *Computers and Operations Research*, vol. 38, no. 1, pp. 320–327, 2011.

[86] F. Legillon, A. Liefooghe, and E. G. Talbi, "A coevolutionary meta-heuristic for bi-level optimization," *Rapport de Récherche INRIA*, vol. 7741, pp. 1–22, 2011.

[87] J. F. Camacho-Vallejo, Á. E. Cordero-Franco, and R. G. González-Ramírez, "Solving the bilevel facility location problem under preferences by a Stackelberg-evolutionary algorithm," *Mathematical Problems in Engineering*, vol. 2014, Article ID 430243, 14 pages, 2014.

[88] B. Huang and N. Liu, "Bilevel programming approach to optimizing a logistic distribution network with balancing requirements," *Transportation Research Record*, no. 1894, pp. 188–197, 2004.

[89] C. M. Feng and C. C. Wen, "A bi-level programming model for allocating private and emergency vehicle flows in seismic disaster areas," in *Proceedings of the Eastern Asia Society for Transportation Studies*, vol. 5, pp. 1408–1423, 2005.

[90] J. Wang, J. Zhu, J. Huang, and M. Zhang, "Multi-level emergency resources location and allocation," in *Proceedings of the IEEE International Conference on Emergency Management and Management Sciences (ICEMMS '10)*, pp. 202–205, August 2010.

[91] H. Sun, Z. Gao, and J. Wu, "A bi-level programming model and solution algorithm for the location of logistics distribution centers," *Applied Mathematical Modelling*, vol. 32, no. 4, pp. 610–616, 2008.

[92] J. Xu and P. Wei, "A bi-level model for location-allocation problem of construction and demolition waste management under fuzzy random environment," *International Journal of Civil Engineering*, vol. 10, no. 1, pp. 1–12, 2012.

[93] P. Hansen, Y. A. Kochetov, and N. Mladenovic, "Lower bounds for the uncapacitated facility location problem with user preferences," G-2004-24, GERAD-HEC, Montreal, Canada, 2004.

[94] I. L. Vasil'ev, K. B. Klimentova, and Y. A. Kochetov, "New lower bounds for the facility location problem with client preferences," *Computational Mathematics and Mathematical Physics*, vol. 49, no. 6, pp. 1055–1066, 2009.

[95] I. L. Vasil'ev and K. B. Klimentova, "The branch and cut method for the facility location problem with clients preferences," *Journal of Applied and Industrial Mathematics*, vol. 4, no. 3, pp. 441–454, 2010.

[96] M. Marić, Z. Stanimirović, and N. Milenković, "Metaheuristic methods for solving the bilevel uncapacitated facility location problem with clients' preferences," in *EURO Mini Conference*, vol. 39 of *Electronic Notes in Discrete Mathematics*, pp. 43–50, Elsevier, 2012.

[97] J. Bhadury, J. H. Jaramillo, and R. Batta, "On the use of genetic algorithms to solve location problems," *Computers & Operations Research*, vol. 29, no. 6, pp. 761–779, 2002.

[98] T. Uno, H. Katagiri, and H. Kato, "An application of particle swarm optimization to bilevel facility location problems with quality of facilities," *Asia Pacific Management Review*, vol. 12, no. 4, pp. 183–189, 2007.

[99] V. Marianov, M. Ríos, and M. Icaza, "Facility location for market capture when users rank facilities by shorter travel and waiting times," *European Journal of Operational Research*, vol. 191, no. 1, pp. 32–44, 2008.

[100] C. M. C. Rodríguez, D. R. S. Peñate, and J. A. M. Pérez, "Competencia espacial por cuotas de mercado: el problema del líder-seguidor mediante programación lineal," *Recta*, vol. 12, no. 1, pp. 69–84, 2011 (Spanish).

[101] H. Kücükaydin, N. Aras, and I. K. Altinel, "Competitive facility location problem with attractiveness adjustment of the follower: a bilevel programming model and its solution," *European Journal of Operational Research*, vol. 208, no. 3, pp. 206–220, 2011.

[102] V. L. Beresnev, "Local search algorithms for the problem of competitive location of enterprises," *Automation and Remote Control*, vol. 73, no. 3, pp. 425–439, 2012.

[103] D. Kress and E. Pesch, "Sequential competitive location on networks," *European Journal of Operational Research*, vol. 217, no. 3, pp. 483–499, 2012.

[104] D. Kress and E. Pesch, "$(r \mid p)$-centroid problems on the networks with vertex and edge demand," *Computers & Operations Research*, vol. 39, no. 12, pp. 2954–2967, 2012.

[105] H. Küçükaydın, N. Aras, and K. Altınel, "A leader-follower game in competitive facility location," *Computers & Operations Research*, vol. 39, no. 2, pp. 437–448, 2012.

[106] M. G. Ashtiani, A. Makui, and R. Ramezanian, "A robust model for a leader-follower competitive facility location problem in a discrete space," *Applied Mathematical Modelling*, vol. 37, no. 1-2, pp. 62–71, 2013.

[107] J. R. Kim, J. U. Lee, and J. Jo, "Hierarchical spanning tree network design with Nash genetic algorithm," *Computers and Industrial Engineering*, vol. 56, no. 3, pp. 1040–1052, 2009.

[108] D. H. Lee, L. Song, and H. Wang, "Bilevel programming model and solutions of berth allocation and quay crane scheduling," in *Proceedings of the 85th Annual Meeting of Transportation Research Board*, Washington, DC, USA, 2006.

[109] P. G. Comba, "Exponentiation cryptosystems on the IBM PC," *IBM Systems Journal*, vol. 29, no. 4, pp. 526–538, 1990.

[110] M. Brown, D. Hankerson, J. Lopez, and A. Menezes, "Software impl ementation of the NIST elliptic curves over prime fields," Research Report CORR 2000-55, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Canada, 2000.

[111] S.-M. Hong, S.-Y. Oh, and H. Yoon, "New modular multiplication algorithms for fast modular exponentiation," in *Advances in Cryptology—EUROCRYPT '96*, vol. 1070 of *Lecture Notes in Computer Science*, pp. 166–177, 1996.

[112] A. Weimerskirch and C. Paar, "Generalizations of the Karatsuba algorithm for efficient implementations," Cryptology ePrint Archive: Report 2006/224, 2006, http://eprint.iacr.org/2006/224.

[113] J. Großschadl, R. M. Avanzi, E. Sava, and S. Tillich, "Energy-efficient software implementation of long integer modular arithmetic," in *Cryptographic Hardware and Embedded Systems—CHES 2005*, vol. 3659 of *Lecture Notes in Computer Science*, pp. 75–90, Springer, Berlin, Germany, 2005.

[114] R. Moreno, J. M. Miret, and F. Sebe, "A hyperelliptic cryptosystem based on the P1363 IEEE standard," in *Proceedings of the International Meeting on Coding Theory And Cryptography (IMCTC '99)*, Medina del Campo, Spain, 1999.

[115] N. P. Smart, "On the performance of hyperelliptic cryptosystems," in *Advances in Cryptology—Eurocrypt'99*, vol. 1592 of *Lecture Notes in Computer Science*, pp. 165–175, Springer, Berlin, Germany, 1999.

[116] J. M. Arroyo and F. D. Galiana, "On the solution of the bilevel programming formulation of the terrorist threat problem," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 789–797, 2005.

[117] C.-C. Liu, J. Jung, G. T. Heydt, V. Vittal, and A. G. Phadke, "The strategic power infrastructure defense (SPID) system. A conceptual design," *IEEE Control Systems Magazine*, vol. 20, no. 4, pp. 40–52, 2000.

[118] A. Roy, D. S. Kim, and K. S. Trivedi, "Cyber security analysis using attack countermeasure trees," in *Proceedings of the 6th Annual Workshop on Cyber Security and Information Intelligence Research*, 2010.

[119] Y.-L. Yuan, Z.-I. Li, and K.-I. Ren, "Quantitative analysis of load redistribution attacks in power systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1731–1738, 2012.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 843976, 14 pages
http://dx.doi.org/10.1155/2014/843976

*Research Article*

# Simulation of Territorial Development Based on Fiscal Policy Tools

**Robert Brumnik,[1] Tamara Klebanova,[2] Lidiya Guryanova,[2] Sergii Kavun,[3] and Olexandr Trydid[4]**

[1] *Metra Engineering Ltd. R&D Group, 1000 Ljubljana, Slovenia*
[2] *Department of Economical Cybernetics, Kharkiv National University of Economics, Kharkiv 61001, Ukraine*
[3] *Department of Information Technologies, Kharkiv Institute of Banking of the University of Banking of the National Bank of Ukraine, Kharkiv 61174, Ukraine*
[4] *Kharkiv Institute of Banking of the University of Banking of the National Bank of Ukraine, Kharkiv 61174, Ukraine*

Correspondence should be addressed to Sergii Kavun; kavserg@gmail.com

Received 11 November 2013; Accepted 22 March 2014; Published 29 May 2014

Academic Editor: Vyacheslav Kalashnikov



Modern approaches to the development of a national economy are often characterized with an imbalanced inflation of some economic branches leading to a disproportional socioeconomic territories development (SETD). Such disproportions, together with other similar factors, frequently result in a lack of economic integrity, various regional crises, and a low rate of the economic and territorial growth. Those disproportions may also conduce to an inadequate degree of the interregional collaboration. This paper proposes the ways of regulating imbalances in the territorial development based upon the fiscal policy tools. The latter can immediately reduce the amplitude of economic cycle fluctuations and provide for a stable development of the economic state system. The same approach is applied to control the processes of transformation of the tax legislation and tax relations, as well as the levying and redistribution of the recollected taxes among the territories' budgets (this approach is also known as a tax policy). To resume, this paper describes comprehensive models of financial regulation of the socioeconomic territorial development that can help in estimating and choosing the right financial policy parameters. These provide the stable rates of the growth of national economies along with a simultaneous decrease in interregional socioeconomic disproportions.

## 1. Introduction

This study will expose the issue of the socioeconomic development of the nonuniformity of regional socioeconomic systems, which is one of the most commonly occurring issues in the world today. Political leaders, economists, and sociologists will often present asymmetry of regional development and growth in the variability of the territory's development, divergence, differentiation, and unbalanced development of some areas and regions. All of these terms are used to describe the situation of dissimilar growth of the separate elements, which forms the socioeconomic system of the region or country and also of the world which lead to negative trends that threaten the integrity of the whole economic system. This situation is the cause of the nonuniformity of development and is regarded as a major factor in the destabilization of, and fundamental changes in, regional policy. The dominant theories of unbalanced development highlight that without government regulation the market objectively increases the regional disparities into a force of cyclicality and market mechanism self-organization, thus establishing the higher development of the economy in some regions and a weakness in others.

Socially oriented strategies of economic space determine development following global priorities for a state's regional policy as a provision of unified social standards and reduction of the interregional economic differentiation. Among the state regulation territorial development tools are, for example, creation of special economic zones, infrastructure development, preparation of the territories for industrial mastering, stimulation for attracting investments, and so forth. This study shows the practice of regional economic management, with the most effective being the financial tools, such as financial support of the regional funds formation, separate economic budget financing developed for sectors within regions, and tax policy. This study presents an existing fiscal (or tax) policy, which highlights some of the redistribution of financial resources, leading to the inevitable infringement of the interests of regional donors, a slowdown in economic growth rates, absence of stimulus for less developed regions, a reduction in subsidization levels, and an increase in competitiveness. Some economic imbalances induce political imbalances and consequently a growth of social intensity, the forming of threats in border regions, an absence of any motivation of interregional collaboration, and a strengthening of centrifugal tendencies. A current situation or event leads to the need for a correction of regional policy, in particular, financial policy. One of the directions for the improvement of regional financial policy includes a model basis development for state financial regulation of the territory's development.

In modern scientific and economic publications, approaches to the simulation of the regional policy are considered [1–15]. In such publications [1, 2] it is considered that such simulation methods of the budget regulation mechanisms are at different hierarchical levels, as well as econometric methods and panel data. The research is presented in [3, 4] and considers some of the questions using the space-lagged models for testing an available overflow effect and the possibility of costs decreasing, which are linked with a stimulation of regional development. In publications [5–10], some of the fundamentals of the socioeconomic system dynamic models and the simulation models of the region are considered. Publications [11–13] consider questions of optimization method development for regional development strategies and analysis of the interregional economic collaboration based on a complex (or a set) of optimization interregional and interbranch models. The works [14, 15] offer a complex (or a set) of models, which give the possibility of determining the state regulation priority of ability to vital spheres of the regional systems and improve the effectiveness of target coordination based upon a multidimensional analysis and adaptive filtration methods.

Enough interest can be recognized in the development of the models to form an effective regional policy. However, some approaches remain unstudied, which allow to obtain an estimate of consistency for tax, budget policy, and dynamics if the investment processes have its influence on the convergence processes of regional development.

The management tool for simulation is based upon the balanced tax policy of the scenario simulation. This allows for research into some of the cause-and-effect factor links, which have an implicit structure. This model also allows the formation of a spectrum of development strategies and implements an estimation of some of the realization consequences for different variants of management actions, which are directed towards eliminating some of the imbalances in the regional systems development [16]. An adequate tool for the scenario realization approach is a simulation model, which provides some experimental possibilities and is linked with an estimation and different management scenario analysis of the region's socioeconomic development. The base conception for a simulation of the financial flows of territories is considered the advantages of the method of system dynamic [17], which give the possibility for an account of all structural relationships among some variables and time aspects of transformations.

In this research, a model development scenario was staged for the financial regulation territorial development, which is directed to form a multicomponent economic system and can provide an emergence of reverse interregional links and sustainable development of the common economy. These scenarios are shown as sequences of some states of the socioeconomic territorial system through the realization of different variants of the financial regional policy.

## 2. Scenario Elaboration Conceptual Scheme for Financial Regulation of the Territorial Development

The proposed conceptual scheme for the scenario elaboration of the financial regulation of the territorial development and the content for the stages of this scheme is shown in Figure 1.

In the first stage, the initial scenario of changes in the characteristics of territorial socioeconomic development (SED), due to the implementation of the adopted fiscal policy, is developed. Meeting the challenges of this phase is carried out using the model alignment imbalances using tax instruments and the proposed simulation model of regional financial regulatory [17]. The financial regulatory model of the territorial development has two main components: (1) the unit of resource allocation and (2) the unit of socioeconomic characteristics of the region.

The purpose of the first block is to simulate the possible value of the regional investment of transfers, subventions, and grants. Targeting of the second block is to simulate the impact of the value of investment transfers, subventions, and grants to the regions for the socioeconomic development of regional systems. Therefore, a simulation model of financial territorial regulation allows conducting multivariate projections of regional and state economic development, depending on the adopted financial regulation state policy. Output data of this stage are inversion scenarios of SETD through accepted tax-budget policy realization.

In the second stage, the formation imbalances analysis in regional development will be performed in the following areas: estimation of the regional socioeconomic development, differentiation of the regional socioeconomic development, estimation of the imbalance of the SED in the regions, and identifying the sources of structural imbalances [18, 19].

Assessment of the level of socioeconomic development is conducted by using the reference object as a method of construction, taxonomic indicator of development. The estimation of the differentiation of SED is a dynamic analysis of regional cluster formations and the analysis of the individual propensity to migrate from regions with low levels of socioeconomic development into a group of regions with

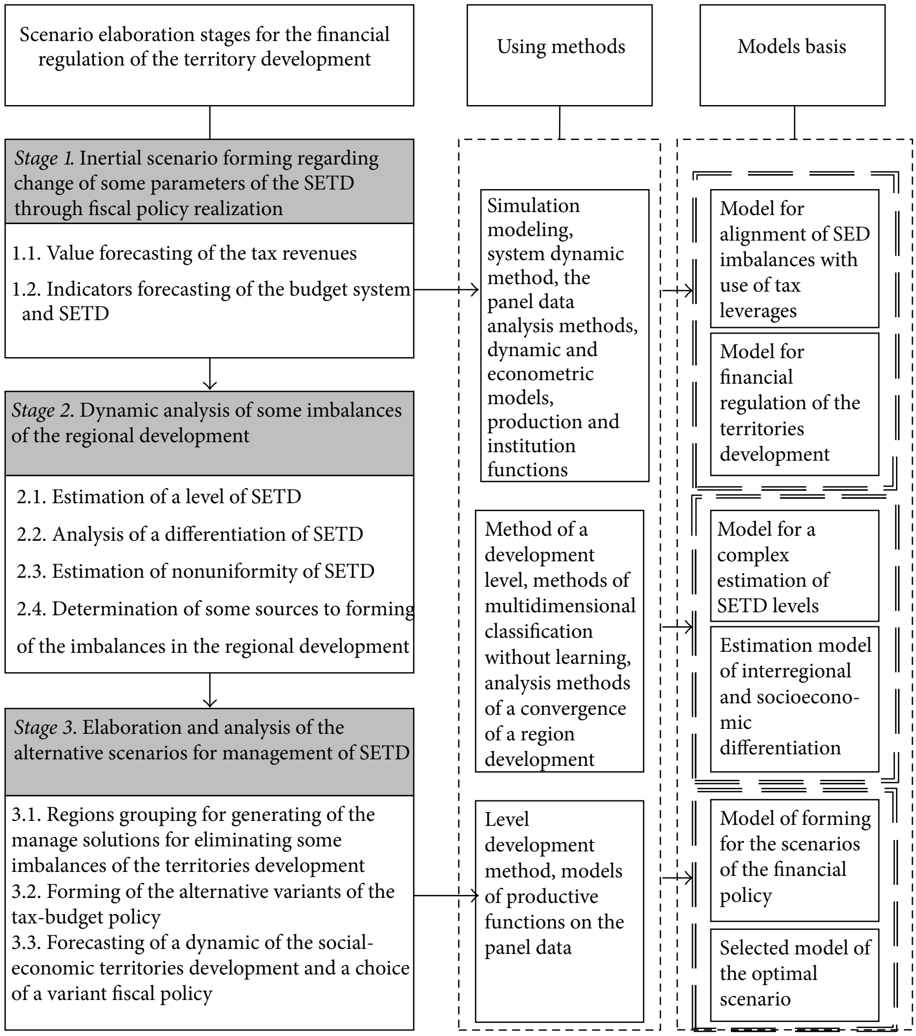



Figure 1: Conceptual scheme of scenario elaboration of the financial regulation of the SETD.

a high level of socioeconomic development. Evaluation is focused on the analysis of the regional structure of the upward or downward trend of economic development. To estimate the irregularity the following data is used: coefficient of variation, coefficient of irregularity (differentiation), coefficient of imbalance, and Tail's index. Identifying the sources of structural imbalances is based upon the decomposition of Tail's index. Herewith, the following factors in increasing the regional imbalances are considered: the unbalanced development of groups of regions with a high level of SED (donor regions) and regions with low SED (recipient regions), the unbalanced development of regions with high levels of socioeconomic development, and the unbalanced development of regions with low levels of socioeconomic development.

At the third stage, the alternative scenarios of managing the development of the territories are formed, aimed at eliminating or preventing the identified structural imbalances whilst maintaining the overall positive trajectory of the national economy development. The objectives of this phase are to generate management decisions concerning the elimination of imbalances in regional development, the formation of alternative fiscal policy options, forecasting the dynamics of socioeconomic development, and selecting an option of fiscal policy. Solving the above tasks is performed by grouping the regions, considering parameters such as the level and rate of socioeconomic development. The following regions and groups were also allocated: *regions leaders* (i.e., regions with a high level and a high rate of socioeconomic development); *stagnant regions* (i.e., regions with the high level and the low rate of socioeconomic development); *developed regions* (i.e., regions with the low level and the high rate of socioeconomic development); and *problematic regions* (i.e., regions with the low level and the low rate of socioeconomic development).

Formation of alternative fiscal policy options suggests changing the parameters of the distribution of investment transfers, in particular the regional development fund among the described groups of the regions. Since the cyclical downturn in the state investment policy is aimed at increasing the speed of investment flow, especially in the production of the high added value, then adjustment of the parameters of the distribution of investment transfers is based upon the research on the territories asset management ratio of the industrial and economic systems (PES) [20]. The predicted results of socioeconomic development of the regions as a result of implementation of the different options of fiscal policy underlie the formation of alternative management scenarios of SETD.

In considering the following scenarios, the alternative compensation scenario assumes an estimation of the aftermaths of the priority investment support of the regions-donors with the implementation of pessimistic scenarios of tax revenues. The main target of the development of this scenario is to evaluate the possibility of forming a "compensatory" effect of reducing the depth of the economic crisis by changing the fiscal policy parameters. The alternative antirecessionary scenario is directed to modeling the results of the phased financial support of the recipient and donor regions. Financial support of the recipient regions allows a reduction in the level of subsidization and in the depth of the economic crisis at the beginning of the implementation of the state stabilization policy. The financial support of the donor regions is aimed at promoting an inward investment in the production of high additional value and preventing the effect of a "deferred" cyclical downturn in the forecast period. Selecting an option of the fiscal policy based on the analysis of the parameters of the regional financial policy offers the alignment of the regional socioeconomic development while maintaining the positive trend in the economy.

Thus, the proposed scheme (Figure 1) of the socioeconomic development scenarios of the regions makes it possible to assess the consistency of fiscal, monetary, and investment policies and to increase the quality of information-analytical basis for decision-making and concerning the elimination of territorial imbalances.

## 3. Financial Regulation Model of the Territorial Development

Developing the scenarios of socioeconomic development of the territories is based on the financial regulation model, which includes simulation models of budget system, indicators, and socioeconomic characteristics of the regions. Together with other similar traditional approaches, such as a method of system dynamics, the proposed simulation models include some other dynamic econometric models, which is a feature of those models. The structurally determined equations are using the panel data models [1], which allow researching the space-dynamic effects of a realization of the fiscal policy.

In general, the panel data model cab will be presented as follows: $y_{it} = \alpha + x'_{it} \times \beta_{it} + u_{it}$, where $y_{it}$ is a value of the researched indicator for $i$th region in $t$th time period, $i = \overline{1,n}$, $t = \overline{1,T}$; $x'_{it}$ is a vector of the explaining variables (factors); $u_{it}$ is a disturbance for the $i$th region (object) in $t$th time period; $\beta_{it}$ is the model parameters (indicators). This model is a general model, so it is suggested that a standard presentation is used for constancy of the parameters (indicators) $\beta_{it}$ for the all values indexed by $t$ and $i$. The panel data specific model can give a possibility for additional division of disturbances on some of the components (parts): $u_{it} = \mu_i + \varepsilon_{it}$, where $\mu_i$ is the unobservable specific and individual effects and $\varepsilon_{it}$ is the residual "noises."

Models with a fixed or random effect depend on suggestions of a property (or character) regarding the value $\mu_i$. If the value $\mu_i$ is $n$th unknown fixed parameters, then this model is one from a set of panel data standard models with fixed effects and it can be presented as follows: $y_{it} = \mu_i + x'_{it} \times \beta + \varepsilon_{it}$. It is suggested that the average level of a dependent variable for the $i$th region can differentiate from an average level of dependent variable $j$th and is also a permanent variable for a different time period and is implemented into this model with the help of different cross-section values $\mu_i$. Different constant values $\mu_i$ will be estimated for different objects. At the same time, these estimated parameters $\beta_i$ would be similar to all the objects and for all time periods. Besides, the general (common) cross-section $\alpha$ will be absent in this case because if it exists, then this cross-section $\alpha$ in this model will raise an effect of multicollinearity.

If it is possible that $\mu_{it}$ is explained as a realization of independent values from $X_{it}$ random variables, then this model belongs to a type of standard panel data model with random effects. In some models with a random cross-section effect, $\mu_{it}$ is defined as a random value, which has a zero expectation, but the perturbations $u_{it}$ are unrelated for different time periods. The model with random effects has the following type: $y_{it} = \alpha + x'_{it} \times \beta + \mu_s + \varepsilon_{it}$, where $\alpha$ is a general cross-section. In these models with random effects, in contrast to those models with a fixed effect, the general cross-section can be identified and separately estimated.

The main contrast among panel data models with fixed and random effects is shown by the cross-section of those models. In the models with a random effect, the cross-sections are considered as random values, whereas the models with a fixed effect—as the fixed values—exhibit differences in some objects.

The model with fixed effects $y_{it} = \mu_i + x'_{it} \times \beta + \varepsilon_{it}$ can be considered as a model with an individual fictitious value; that is, for each region, a variable is introduced which has an individual character. Implying the existence of the same parameters for all regions for all time moments, the availability of heterogeneity among some observation objects with the invariant objects with respect to time, but with a specific location parameter for each object, can be researched. By introducing fictitious variables,

$$d_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases} \tag{1}$$

for each object, this model can be presented as a standard model: $y_{it} = \sum_{j=1}^{n} \mu_j d_{ij} + x'_{it}\beta + \varepsilon_{it}$. For an estimation of these parameters the intragroup transformation is used—$y_{it} - \overline{y}_i = (x_{it} - \overline{x}_i)'\beta + \varepsilon_{it} - \overline{\varepsilon}_i$. If the ordinary least squares- (OLS-) method is used, leading to a regression through the beginning of the origin of coordinates, a consistent estimate will be obtained:

$$\widehat{\beta}_{\mathrm{FE}} = \left( \sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)' \right)^{-1} \cdot \sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(y_{it} - \overline{y}_i). \tag{2}$$

Due to the subtraction of the time average of the data, this estimation is accountable with a variation in the borders of the object of observation. The estimations of individual effects can be used as $\mu_i = \acute{y}_i - \acute{x}_i' \cdot \widehat{\beta}_{\mathrm{FE}}$. These estimations are unbiased and consistent estimations for fixed $n$ at $t \rightarrow \infty$. A formula for a covariance matrix has the following type: $V(\widehat{\beta}_{\mathrm{FE}}) = \sigma_{\varepsilon}^2 (\sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)')^{-1}$. As a dispersion estimation, $\sigma_{\varepsilon}^2$ can use $\widehat{\sigma}_{\varepsilon}^2 = (1/(nT - n - k)) \sum_{i=1}^{n} \sum_{t=1}^{T} (y_{it} - \overline{y}_i - (x_{it} - \overline{x}_i)' \widehat{\beta}_{\mathrm{FE}})^2$. Here a bilevel procedure was used, whereby at first a calculated estimation was made with an account of the variation among the objects of observation and, afterwards, a calculation of the individual effects is provided.

The model with random effects cannot effectively be estimated either, with the help of the OLS-method, because some mistakes in suggestions for this model are correlated between them due to the presence of a specific summa for each object of observation. Therefore, the bisteps procedure FGLS is used:

$$\widehat{\beta}_{\mathrm{RE}} = W \widehat{\beta}_{\mathrm{FE}} + (I_k - W) \widehat{\beta}_b, \tag{3}$$

where

$$\widehat{\beta}_B = \left( \sum_{i=1}^{n} T(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})' \right)^{-1} \sum_{i=1}^{n} T(\overline{x}_i - \overline{x})(\overline{y}_i - \overline{y}),$$
$$W = \left( S_{xx}^{w} + (1-\theta)^2 S_{xx}^{b} \right)^{-1} S_{xx}^{w}, \tag{4}$$
$$S_{xx}^{w} = \sum_{i=1}^{n} \sum_{t=1}^{T} (x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)',$$

where $I_k$ is the identity matrix with dimension $k$ (the explanatory variables). The parameter $\theta$ is selected as follows in order that the mistakes in the model will not be interrelated at any time for different values $t$. Then the parameter $\theta$ assumes the following form:

$$\theta = 1 - \frac{\widehat{\sigma}_{\varepsilon}}{\sqrt{\widehat{\sigma}_{\varepsilon}^2 + T \widehat{\sigma}_M^2}}; \qquad \widehat{\sigma}_M^2 = \frac{\sum_{i=1}^{n} u_i^2}{n-k-1} - \frac{1}{T} \widehat{\sigma}_{\varepsilon}^2, \tag{5}$$

where $u_i$ is the remains, which received with the help of the OLS-method in a regression with accounting a variation among objects of observation $\overline{y}_i = \mu + \overline{x}_i' \beta + u_i + \overline{\varepsilon}_i$, $i = \overline{1, n}$.

It should be noted that one of the problems of using panel data is that of choosing a model type (usual regression, fixed, or random effect). In the above models, a definite hierarchy exists. The usual regression model is a special case of the model with fixed effects, when in this model we have $\mu_i = 0$, $i = \overline{1, n}$. In addition, this model can be used as the model with a random effect (when the mistakes are absent) or the model with fixed effects (when a correlation between $\mu_{it}$ and $X_{it}$ is absent). Therefore, with statistical tests using a null hypothesis, there is a possibility of using a particular model, but the use of alternative tests of a null hypothesis means the possibility to use a common (general) model. The choice of model specification exists and is based on the $F$-test, Breusch-Pagan's test, and Haussmann's test.

*F-test*. A null hypothesis is formed to check the statistical significance of cross parameters in a model $H_0$ where $H_0$: $\mu_i = \mu_j$ for any $i$, $j$. It corresponds to a model with the same parameter $\mu$ for all objects sampled, that is, to an aggregated model. The alternative hypothesis consists of the following fact that $H_1$: $\mu_i \neq \mu_j$, at least for one pair $i$, includes of a model with fixed effects. The calculated criteria value is calculated as follows:

$$F = \frac{R_{\mathrm{pool}}^2 - R_{\mathrm{FE}}^2}{1 - R_{\mathrm{FE}}^2} \frac{nT - n - k}{n - 1} = \frac{Q_{\mathrm{pool}} - Q_{\mathrm{FE}}}{Q_{\mathrm{FE}}} \cdot \frac{nT - n - k}{n - 1} \overset{H_0}{\sim} F(n-1, nT - n - k). \tag{6}$$

*Breusch-Pagan's Test*. In order to check the statistical significance of the random effects, a calculation is made with the help of Lagrange's multiplier test:

$$\mathrm{LM} = \frac{nT}{2(T-1)} \left( \frac{\sum_{i=1}^{n} \left( \sum_{t=1}^{T} e_{it} \right)^2}{\sum_{i=1}^{n} \sum_{t=1}^{T} e_{it}^2} - 1 \right)^2, \tag{7}$$

where $e_{it}$ is the oddments of the aggregated regression model.

If the hypothesis $H_0$ is true, then the statistic LM has $\chi^2$ distribution and is valid with a single degree of freedom.

*Haussmann's Test*. Due to the most important difference in the heterogeneity simulation approaches of observable objects, there is a relation among input effects with some regressors (undefined variables). In this case, the random effects are suggested as interrelated with these regressors, whereas the fixed effects can be correlated with them. In addition, the choice of model with fixed or random effects depends on whether these effects are correlated or not with those regressors. During the validity of the hypothesis $H_1$, the model estimations with the fixed effects exist, but the model estimations with the random effects do not exist. In this case, some significant differences can be found between estimations of these two models. The value of Haussmann's statistic is calculated as follows:

$$H = \left( \widehat{\beta}_{\mathrm{FE}} - \widehat{\beta}_{\mathrm{RE}} \right)' \widehat{\Phi}^{-1} \left( \widehat{\beta}_{\mathrm{FE}} - \widehat{\beta}_{\mathrm{RE}} \right), \tag{8}$$

where the estimation $\widehat{\Phi}$ is for the covariance matrix $(D(\widehat{\beta}_{\mathrm{FE}}) - D(\widehat{\beta}_{\mathrm{RE}}))$.

Such modelling is based upon the consideration of information and data for building models of a budget system and data on the socioeconomic development of Ukraine in the last 15 years.

A simulation model of the dynamics of the budget system indicators, which includes such variables as revenues of the consolidated budget, health care expenses, educational expenses, social protection and social security expenses, economic activity expenses, state budget expenses, donations, grants, and investment transfers to the regions, is based on a system of 11 equations. Some of these are presented below.

*Health Care Expenses*

$$\begin{aligned}\text{Zatr\_oh\_zd}_t = {} & 0.06704 \cdot \text{Dohod\_svodnogo\_budgeta}_t \\ & + 0.021184 \cdot \text{VVP}_{t-1},\end{aligned} \tag{9}$$

where $\text{Zatr\_oh\_zd}_t$ is the health care expenses (million UAH, is the ISO 4217 currency code of the Ukrainian hryvnia, the national currency of Ukraine), $\text{Dohod\_svodnogo\_budgeta}_t$ (is the variable, which describes the revenue (income or profit) part of the aggregate budgets: operational, investment, and financial) is the revenues of the consolidated budget (million UAH), and $\text{VVP}_{t-1}$ are the gross domestic product (GDP), with a lag equaling 1 (million UAH).

The determination coefficient of the dynamical model was 0,99573; the following calculated values of $t$-test 3,0658 and 2,75002 are greater than the table value, which allows the conclusion that estimates of the model parameters are statistically significant; the value of the Durbin-Watson statistic equal to 1,56 indicates the absence of autocorrelation in a number of residuals.

*Educational Expense*

$$\begin{aligned}\text{Zatr\_obbrazov}_t = {} & -4853.6 + 0.12 \\ & \cdot \text{Dohod\_svodnogo\_budgeta}_t \\ & + 0.04 \cdot \text{VVP}_{t-1},\end{aligned} \tag{10}$$

where $\text{Zatr\_obbrazov}_t$ are the educational expenses (million UAH).

The determination coefficient of the dynamical model was 0,994; calculated values of $t$-test, 3,07, 3,56, and 3,88, are greater than the table value at a 1% level of significance, which leads to the conclusion that estimates of the model parameters are statistically significant.

The general (common) view of a simulation model of the budget system indicators is shown in Figure 2.

In the socioeconomic simulation model characteristics of the territory development, such variables are considered as gross regional product, total export volume, investment in fixed assets, level of employment, the total import volume, volume of innovative products, value of foreign investments, the average monthly wage, income, level of economically active population, and the provision of housing, bringing into service the apartments and the number of students at universities. A few equations are presented in Figure 2.

*Gross Regional Product*

$$\begin{aligned}\text{VRP}_{it} = {} & \mu_i^1 + 2474.059 \cdot \text{Export}_{it} + 0.566316 \cdot \text{IOK}_{it} \\ & + 62.60181 \cdot \text{Zan\_nas}_{it} + 0.888341 \cdot \text{VRP}_{it-1},\end{aligned} \tag{11}$$

where$\text{VRP}_{it}$ is the gross regional product for the $i$th region per capita (UAH); is the fixed effect in the $i$th region; $\text{Export}_{it}$ is the total export volume of the $i$th region per capita (thousand USD); $\text{IOK}_{it}$ is the investments in fixed assets of the $i$th region per capita (UAH); $\text{Zan\_nas}_{it}$ is the level of employment of the $i$th region (% of population aged 15–70 years); $\text{VRP}_{it-1}$ is the gross regional product for the $i$th region per capita with lag equaling 1 (UAH).

The coefficient of considered panel data model determination was 0,978; calculated values of $t$-test, 48,448; 2,899; 7,86; and 8,56, are greater than the table value, which leads to the conclusion that estimates of the model parameters are statistically significant.

*The Total Export Volume*

$$\begin{aligned}\text{Export}_{it} = {} & \mu_i^2 + 0.434655 \cdot \text{Export}_{it-1} + 0.132915 \cdot \text{Export}_{it-2} \\ & + 0.0000054 \cdot \text{VRP}_{it-1} + 0.0000591 \cdot \text{OB\_IP}_{it},\end{aligned} \tag{12}$$

where the fixed effect is in the $i$th region; $\text{OB\_IP}_{it}$ is the volume of innovative products of the $i$th region per capita (UAH); $\text{Export}_{it-1}$ is the total export volume of the $i$th region per capita with lag equaling 1 (thousand USD); $\text{Export}_{it-2}$ is the total export volume of the $i$th region per capita with lag equaling 2 (thousand USD).

The determination coefficient of the considered panel data model was 0,854491; calculated values of $t$-test, 2,7167; 2,7494; 6,02; and 1,77, are greater than the table value, which leads to the conclusion that estimates of the model parameters are statistically significant.

*The Total Import Volume*

$$\begin{aligned}\text{Import}_{it} = {} & \mu_i^3 + 0.4263 \cdot \text{Import}_{it-1} + 0.0000886 \cdot \text{OB\_IP}_{it} \\ & - 0.0000499 \cdot \text{OB\_IP}_{it-1} + 0.0000054 \cdot \text{VRP}_{it-1},\end{aligned} \tag{13}$$

where the fixed effect is in the $i$th region; $\text{Import}_{it}$ is the total import volume of the $i$th region per capita (thousand USD); $\text{OB\_IP}_{it-1}$ is the volume of innovative products of the $i$th region per capita with lag equaling 1 (UAH); $\text{Import}_{it-1}$ is the total import volume of the $i$th region per capita with lag equaling 1 (thousand USD).

The determination coefficient of the considered panel data model was 0,7117; calculated values of $t$-test, 5,81; −2,39; 4,58; and 4,69, are greater than the table value, which leads

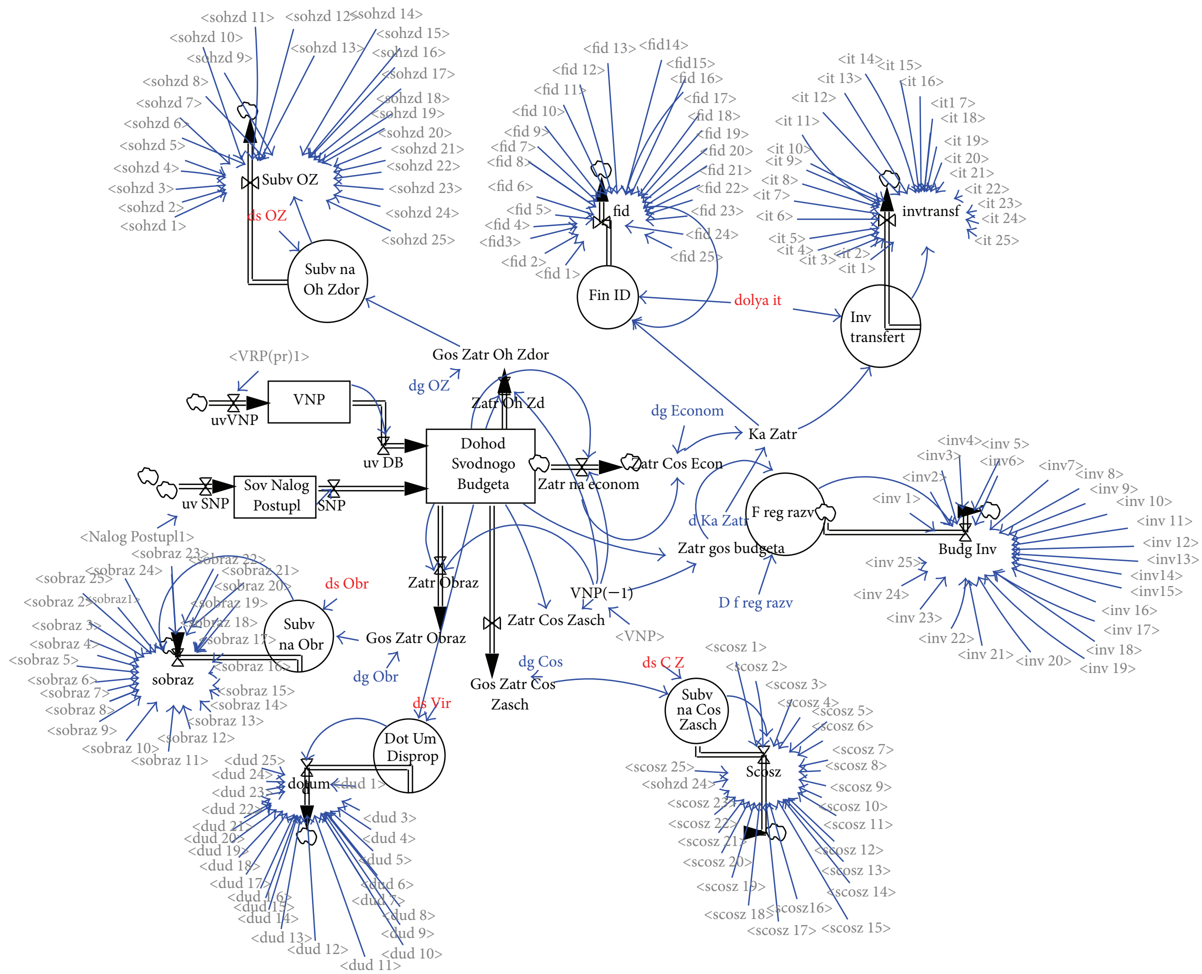



Figure 2: Simulation model for budget system indicators.

to the conclusion that estimates of the model parameters are statistically significant.

*The Average Monthly Wage*

$$\text{Sr\_ZP}_{it} = \mu_i^4 + 22.44544 \cdot \text{Zan\_nas}_{i,t-1} + 0.097669 \cdot \text{VRP}_{it}, \quad (14)$$

where $\text{Sr\_ZP}_{it}$ is the average monthly wage in the $i$th region (UAH); is the volume for a fixed effect for the $i$th region; $\text{Zan\_nas}_{it-1}$ is the level of employment in the $i$th region with lag equaling 1 (% of population aged 15–70 years).

The general (common) type of the simulation model of a dynamic of the socioeconomic region indicators is shown in Figure 3.

The determination coefficient of the considered panel data model was 0,9727; calculated values of $t$-test, 5,39 and 56,609, are greater than the table value, which leads to the conclusion that estimates of the model parameters are statistically significant.

The list of exogenous simulation model parameters sets is presented in Table 1.

The comprehensive country financial regulatory model of the socioeconomic region development includes a simulation model of the budget system indicators and 25 simulation models of the socioeconomic region indicators (characteristics).

With analysis of some results, which were received by these simulation models, it is possible to make a highly precise forecast of economic value and confirm the necessity for use of scenario development. The comparison results for the actual and calculated values of the socioeconomic regional development studied indicators are shown in Figures 4(a)–4(d). Other indicators received similar results.

Analysis of the findings leads to the conclusion that the simulation model provides the highest accuracy in forecasting such indicators as GDP per capita, the interest rate on the loan, the level of employment, and the provision of housing. The prediction error for these parameters varies from 0,412% to 8,8649%. Similar results were also obtained for the other indicators.

Generally, the prediction error of the simulation model of financial regulation for the regional development does not exceed 11%, which proves the possibility of its utilization for the socioeconomic development of the territories of the characteristic change scenarios due to the implementation of different fiscal policy variants.

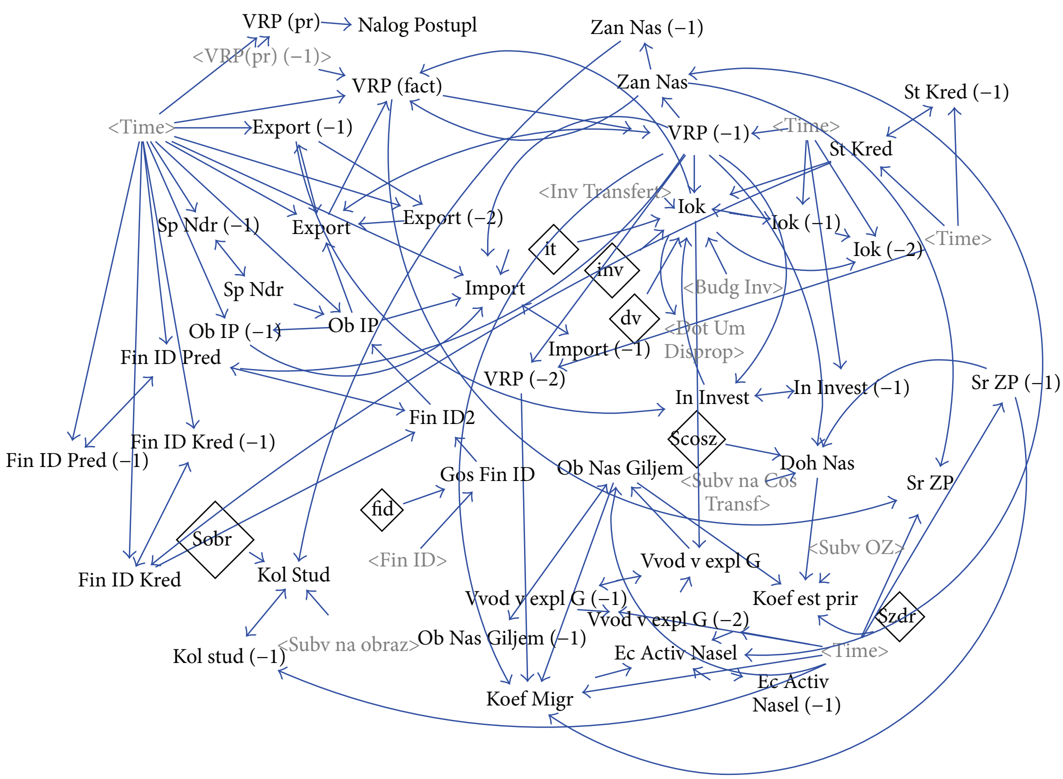



Figure 3: Dynamic simulation model of the socioeconomic region indicators.

Table 1: The values of exogenous parameters set.

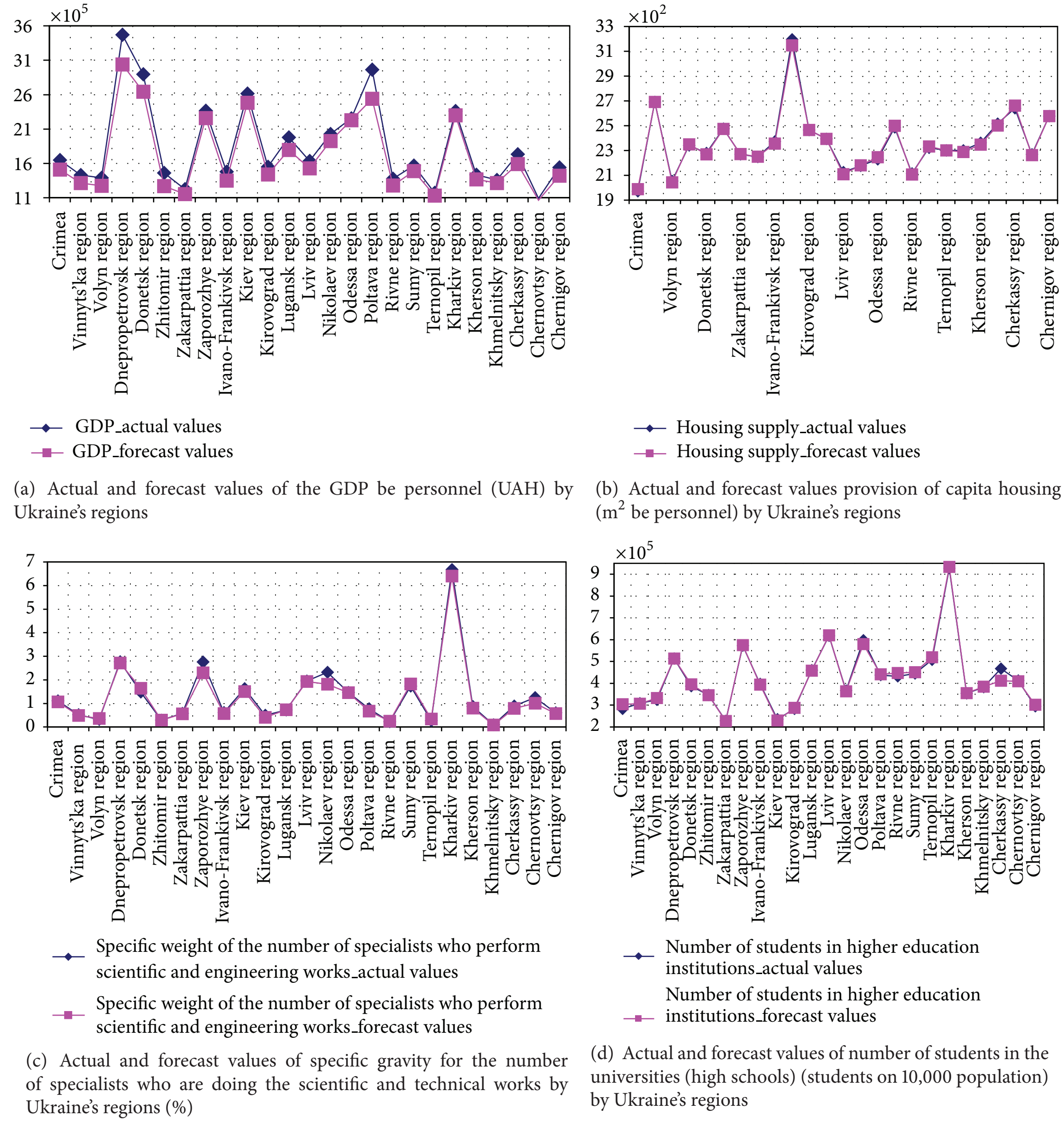
| Symbol | $\mathrm{Szdr}_i$, $i=\overline{1,25}$ | $\mathrm{Sobr}_i$, $i=\overline{1,25}$ | $\mathrm{Scosz}_i$, $i=\overline{1,25}$ | $It_i$, $i=\overline{1,25}$ | $\mathrm{Inv}_i$, $i=\overline{1,25}$ | $D\_V_i$, $i=\overline{1,25}$ | $\mathrm{Fid}_i$, $i=\overline{1,25}$ |
|---|---|---|---|---|---|---|---|
| | | Coefficients of the subventions distribution | | | | Coefficients distribution | |
| Description | For health | For education | For the social protection and security | For the investment transfers | Public investment | Equalization grants | Funding innovation |

## 4. Forming of the Inertial Scenarios and Analysis of the Imbalances for the Regional Development

According to the demonstration in the scheme in Figure 1 at the first stage of the formation, the inertial scenarios were performed, which assumes the study of the dynamic changes of the characteristics of the socioeconomic development territories due to the adopted fiscal policy implementation. In the scenario development, an account was made of tax reforms, regarding the reduction of a tax burden on legal persons, reduction of value-added tax, an increase in the tax burden on physical persons with a high-level income, and an increase in the excise tax for such goods groups as alcohol, tobacco, fuel, and cars to increase of the import customs duties.

Calculating the value for the interbudget transfer changes in the budget policy, with respect to forming of a regional development fund, was taken into account. Also, two inertial scenarios were developed: (1) the optimistic inertial scenario, which is accounted for as an extension of the tax base owing to a change of parameters of the tax policy, and (2) the pessimistic inertial scenario, which is based upon the availability of a time lag indicating the positive “feedback” in an economy and, as a result, informing of the budget deficiency.

An estimation of the socioeconomic regional development level dictates a realization of the accepted fiscal policy, which existed with the help of the development level method [8]. The choice of this method is due to the following advantages: absences of some limits for a number of signs for the information space (a number of indicators can be included in an indicator system, a positive dynamic of which is confirmed by a reduction in the growth of the SER level); an original indicator system can include some signs which have a different dimension, some values of an integral indicator which have a normative range of changes that provides an interpretation of the received results. The level development indicator is calculated by the formula $d_i^* = 1-(c_{i0}/c_0)$, where $c_0 = \bar{c}_0 + 2S_0$, $\bar{c}_0 = (1/n)\sum_{i=1}^{n} c_{i0}$, $S_0 = (1/n)\sqrt{\sum_{i=1}^{n}(c_{i0}-\bar{c}_0)^2}$, and $c_{io} = \sqrt{\sum_{j=1}^{m}(Z_{ij}-Z_{0j})^2}$ is the Euclidean distance among points-units (regions) and the point $P_0\ (z_{01}, z_{02}, \ldots, z_{0m})$, which is a developed etalon. Fundamentals for the building of




(a) Actual and forecast values of the GDP be personnel (UAH) by Ukraine's regions

(b) Actual and forecast values provision of capita housing ($m^2$ be personnel) by Ukraine's regions

(c) Actual and forecast values of specific gravity for the number of specialists who are doing the scientific and technical works by Ukraine's regions (%)

(d) Actual and forecast values of number of students in the universities (high schools) (students on 10,000 population) by Ukraine's regions

Figure 4: Compared actual and forecasted values of the indicators of the socioeconomic regional development.

a developed etalon are using a sign separation on a stimulant and destimulant. Some signs, which provide a positive and stimulating influence on the level of the socioeconomic regional development, are called stimulants, in contrast to the sign-destimulants. The coordinates of the developed etalon are calculated as follows:

$$z_{0j} = \begin{cases} \max\limits_i z_{ij}, & \text{если } j \in I, \\ \min\limits_i z_{ij}, & \text{если } j \notin I, \end{cases} \tag{15}$$

where $I$ is the set of stimulants. Due to the sign-destimulants having different dimensions, then at the formation of the matrix of the distances $C = (d_{i0})$, $i = \overline{1,n}$, is existing their standardization based on the following formulas: $z_{ij} = (x_{ij} - \overline{x}_i)/S_j$, $S_j = \sqrt{\sum_{i=1}^{n}(x_{ij} - \overline{x}_j)^2/(n-1)}$. The values of the integral indicator change in the range from 0 to 1. The nearer values to 1 of this indicator mean the higher level of socioeconomic region development.

The results of the integral indicator calculation on the level of the socioeconomic region development, received and based on average values of socioeconomic development indicators for 25 regions in the forecast period and characterized by a tendency for the development of a national economy in general, are shown in Figure 5.

From Figure 5, the stabilization policy is available, and positive effects are obtained in the medium perspective by two developed scenarios. This fact is confirmed by a dynamic value for the integral indicator for 2013-2014. Due to the reduction of integral indicator values in 2015 the possibility

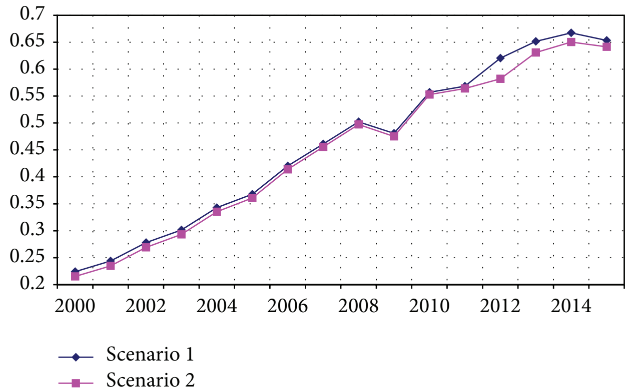



Figure 5: Integral indicator dynamic on the level of socioeconomic regional development.

of creating a conclusion about the forming of a descending developed tendency and occurrence of the "differed" cyclical downturn was obtained.

The economic space proportions were studied with the help of the cluster analysis methods [8, 9, 16]. Classification consists of the decomposition of an initial set of regional points to make a comparison of a small number of classes $Q = \{Q_1, Q_2, \ldots, Q_n\}$. As for the regions which are owned by one class (or group), these would be placed at comparatively small distances from each other.

The established similarity or differences among regions (or among classified objects) are dependent on the metric distance between them. The following distances are used to measure among objects (Table 2).

The final conclusions confirm the necessity of a distribution parameter correction for interbudget transfers with the purpose of saving positive developed tendencies of a national economy. In this connection, a forecasted dynamic analysis was made of the socioeconomic region development during 2013–2015. An analysis was made of the structure of the region's cluster formations, a tendency of separated regions of migration of a group with a low level of socioeconomic development to a group with a high level of SETD.

For the building of the group of the hierarchical agglomerative and iterative methods were used. The hierarchical agglomerative methods can only give a conditionally optimal solution in a subset of a local portion (or cluster). However, an advantage of these methods is the calculated simplicity and interpretation of the results received. An entity of the hierarchical agglomerative method is included in the fact that, on the first step, each selected object is considered as a separated cluster. The combining clustering process occurs successively; the most similar objects unite which are based on the distance matrix or the similarity matrix. Clusterization results, which are presented as a dendrogram, allow the presentation of a hypothesis about the number of clusters. This number of clusters is used for choosing initial conditions for the iteration algorithm, which is based on the method of "$k$-means" [14]. The algorithm of this method includes the following fundamental steps: $k$ points (or regions) are randomly selected or are indicated from $n$ regions by a researcher based on some prior considerations.

These points are like etalons; a sequence number, which at the same time is a cluster number, is assigned to each etalon; from the remaining ($n$-$k$) regions the point $X_i$ is retrieved with the coordinates $(x_{i1}, \ldots, x_{im})$ and it is checked for which etalons (or centers) it is closest to. The checked region joins that center (or etalon), to which $d_l = \min d_{il}$ $(l = 1, \ldots, k)$ matches. This etalon is replaced by the new one, which was recalculated with an account of a jointed point, and its weight (a number of regions are included in this cluster) is increased by one. If there are two or more minimal distances, then $i$th region is attached to a center with the smallest sequence number; afterwards, the point $X_{i+1}$ will be selected and all the steps (procedures) are repeated.

Thus, via ($n$-$k$) steps, all points (regions) of a set will be assimilated to one of the $k$ clusters, but the partition process is not finished with this step. In order for the stability of a partition to be received according to a similar rule, then all points $X_1, X_2, \ldots, X_n$ are connected to received clusters again, whilst at the same time the weights continue to accumulate. The new partition is compared with a previous portion and they are the same; then this algorithm is completed. In another case, this cycle is repeated. The final partition may have some gravity centers, which do not match with the etalons, so they can be classified as $C_1, C_2, \ldots, C_n$. At the same time, each point $X_i$ $(i = 1, 2, \ldots, n)$ will relate to that cluster (or class), for which the following $D(x_j, c_l) = \min d(x_j, c_j)$ pertains.

After finishing the classification procedures, there is a need to estimate the received results. For this, a measure of a classification quality can be used, a so-called "the functional of a quality." The best partition should be considered as such a partition by selecting the functional, of which an extreme value of the objective function, the functional of a quality, can be achieved. The following functionals of a quality were considered in the partition analysis: $F_1 = \sum_{l=1}^{k} \sum_{i \in S_l} d^2(x_i, \overline{x}_l)$, $F_2 = \sum_{l=1}^{k} \sum_{i,j \in S_l} d_{ij}^2$, $F_3 = \sum_{l=1}^{k} \sum_{j=1}^{p} \sigma_{lj}^2$. Optimal partition is a partition where we have $F_i \rightarrow \min_{S \in A}$ and where $A$ is the set of all allowable partitions.

For determining a cluster number, which is needed to split the original region set, the hierarchical agglomerative methods were used. As a distance, the Euclidean distance measure was considered. For determining the distances among a random cluster pair, Ward's method was used, which allowed the minimization of the sum of squared deviations among each region (object) and a cluster center, which was included in this region. A graphical analysis of the grouped results, which were presented as a dendrogram (Figure 6), allowed the conclusion to be made that the researched regions set could be split into two objects, or groups, similar to the socioeconomic developmental characteristics.

The method of "$k$-means" was used as a stable analysis for receiving classification. The results obtained allowed a conclusion to be made regarding the saving of a socioeconomic territories development differentiation in a forecast period. Region clusters with a high and low level of socioeconomic development are sustained by their own structure.

TABLE 2: Distances measures among objects (regions).

| Distances measure among objects | Calculated formula | Conditions of using |
|---|---|---|
| The common Mahalanobis distance | $d_{ij} = \sqrt{\left(X_i - X_j\right)' \Lambda' S^{-1} \Lambda \left(X_i - X_j\right)}$, where $d_{ij}$ is the distance between $i$th and $j$th regions (or objects); $X_i$, $X_j$ are the value vectors of the SETD indicators for $i$th and $j$th objects; $S$ is the common covariation matrix; $\Lambda$ is the matrix of weighting coefficients | Is used in case of dependent vector components $x_1, x_2, \ldots, x_n$ and their different significance at classification |
| Euclidean's distance | $d_{ij} = \sqrt{\sum_{k=1}^{m} \left(x_{ik} - x_{jk}\right)^2}$, where $x_{ik}$, $x_{jk}$ are the values for $k$th indicator, respectively, for $i$th and $j$th objects | Is used in case if the vector observation components $X$ are homogeneous by physical sense and are equally important for classification |
| Weighted Euclidean's distance | $d_{ij} = \sqrt{\sum_{k=1}^{m} w_k \left(x_{ik} - x_{jk}\right)^2}$, where $w_k$ is the weight, which can be used for $k$th indicator | Is used in case when for each vector component $X$ can be used as a weight, which is proportional for a degree of a sign important $0 \leq w_k \leq 1$ |
| Hamming's distance | $d_{ij} = \sum_{k=1}^{m} \left\lvert x_{ik} - x_{jk} \right\rvert$ | Is used as a difference measure of the objects, which can be defined by dichotomous signs |

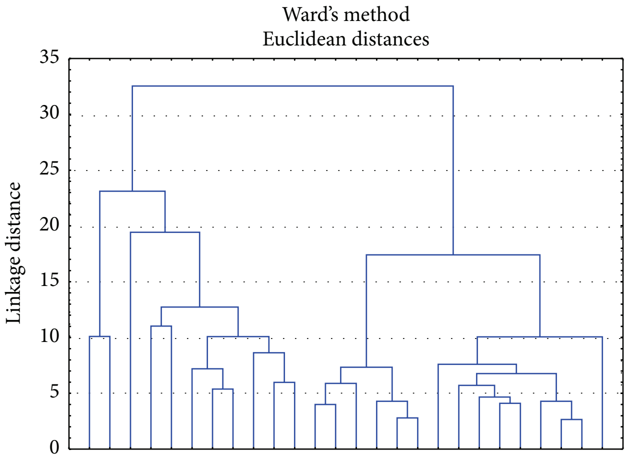



FIGURE 6: Classification dendrogram.

For example, 15 regions, which are included in a group with low developed levels in the preforecast classification and in the forecast period, are saving their own position. A tendency to migrate from one cluster to another is observed in just two regions. A specific regional weight, with a high and low level of SETD, is shown in Table 3.

The results obtained allow the conclusion to be made that in the forecast period a convergence in regional development levels was observed. A value of one of the nonuniform indicators (variation coefficient) for the end of the forecast period in the first development scenario is 18,043% and in the second development scenario is 22,707%. Consequently, we can explain the decrease in the financial possibilities for an uneven smoothing of the territorial development. However, some uneven indicators, which were received from the end of the forecast period, are confirmed as a decrease in interregional differentiation on account of allowed regional policy.

The final direction for an analysis of the socioeconomic territorial development is the detection of some "factors-sources" of structural imbalances, which are based on a decomposition of Tail's index. At the same time, the following amplification factors of the regional imbalances are considered: the imbalance of a development among the groups by region-donors and regional-recipients and imbalanced development into a region's group with a high and low level of socioeconomic development.

As is shown by the analysis that was carried out, one of the general factors which formed was from an imbalance based upon intergroup socioeconomic differentiation. However, some minor decreases in the forecast period of this differentiation can be seen. Unfortunately, some divergent processes are characterized by a group of regional-donors; this is a source of the occurrence for an effect of "deferred" cyclical downturn (shown in Figure 5). Rates comparable to the economic region's growth, with high and low developed levels in the preforecast and forecast periods, show that the preforecast period of outstripping rates was characterized for the first regional group, when the forecast period of the economic growth rates for this group is significantly slowed down.

Therefore, the regions-donors are one of the primary sources for the formation of regional imbalances; they are also confirming the necessity for a correction of the tax-budget policy.

## 5. Forming and Analysis of Alternative Scenarios of Socioeconomic Regional Development

In forming the alternative scenarios of regional policy, the regions were grouped by level and rated according to the socioeconomic region development. The research regions, selected by classification variables, will allow the following

Table 3: Specific regional weight with a high and low level of SETD.

| Specific region weight | Preforecast classification | The period of feed forward 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| With the low level of SETD | 64% | 72% | 64% | 64% | 64% |
| With the high level of SETD | 36% | 28% | 36% | 36% | 36% |

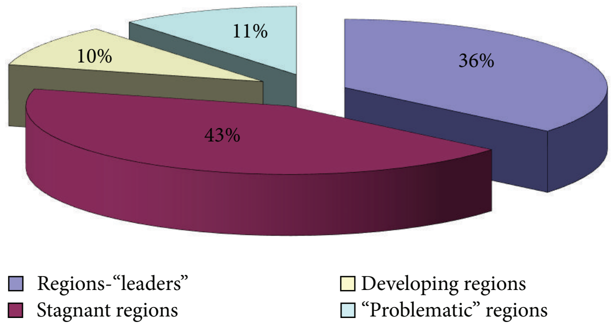



Figure 7: The specific weight of an investment transfer, which was selected from the region's groups.

groups: region-"leaders" (regions with a high level and rate of SETD); stagnant regions (regions with a high level and low rate of SETD); developing regions (regions with a low level and high rate of SETD), and "problematic" regions (regions with the lowest level and a low rate of SETD). An analysis of the specific weight of investment transfers, which was selected from these groups of regions (as is shown in Figure 7), allows the conclusion that an accepted financial and regional policy is directed, at least, to supporting the stagnant and problematic regions.

It can be concluded that the accepted policy levels, for alignment of socioeconomic regional development, lead on one side to an imbalance and decrease in regional development on account of a level reduction of regional depression and accelerated regional growth with a low level of socioeconomic regional development. However, it leads to a significant slowdown in the rate of growth for a region's group with a high level of socioeconomic regional development.

As a result of the above, the priority of the country-wide investment policy, under some conditions of cyclical downturn, is an increase in the investment and innovation activity of the enterprises, which are making products with high added value. A parameter correction is necessary if the fiscal policy is based on analysis of the impacts of resources in regional systems. Research methods, which analyze the impact of resources, are production functions, based on panel data without accounting for a factor of scientific progress. The variants of the panel data model were considered and are defined as

$$\ln \mathrm{VDS}_{it} = \ln \beta_{0i} + \rho_i t + \beta_{1i} \ln \mathrm{Zan}_{it} + \beta_{2i} \ln \mathrm{IOK}_{it} + \varepsilon_{it},$$
$$\ln \mathrm{VDS}_{it} = \ln \beta_{0i} + \beta_{1i} \ln \mathrm{Zan}_{it} + \beta_{2i} \ln \mathrm{IOK}_{it} + \varepsilon_{it}, \quad (16)$$

where $\mathrm{VDS}_{it}$ is the gross value added per capita (in UAH) for $i$th region in $t$th time period:

$\mathrm{Zan}_{it}$is the employed population (thousand populations per one thousand people) for $i$th region in $i$th time period;

$\mathrm{IOK}_{it}$ is the rate of investment in fixed capital per capita (in UAH) for $i$th region in $i$th time period;

$\varepsilon_{it}$ is the random component; $\beta_{0i}$, $\beta_{1i}$, $\beta_{2i}$, and $\rho_i$ are unknown parameters, which need to be quantitatively estimated.

Moreover, some hypotheses were tested for the fact of separated regions with the parameter estimations $\beta_{0i}$, $\beta_{1i}$, $\beta_{2i}$, and $\rho_i$ having no significant regional differences, that is, $\beta_{o1} = \beta_{02} = \cdots = \beta_{025} = \beta_0$; $\beta_{11} = \beta_{12} = \cdots = \beta_{1,25} = \beta_1$; $\beta_{21} = \beta_{22} = \cdots = \beta_{2,25} = \beta_2$; $\rho_1 = \rho_2 = \cdots = \rho_{25} = \rho$.

Due to the scientific progress being differentiated by some types of economic activity (in particular, the higher rates are observed in such industry sectors as mechanical engineering, the communications industry, and instrument making), research and detection were carried out on some types of economic activity that were characterized using the most effective technologies, which formed a "regional profile" of the impact of industry resources.

Similarly, as with the model variants presented above, which account for some regional differences in investment activity effectiveness for a dependence specification, accounting for some industry differences, the model variance was considered and is defined as

$$\begin{aligned} \ln \mathrm{VDS(ED)}_{it} &= \ln \beta_{0i} + \rho_i t + \beta_{1i} \ln \mathrm{Zan}\,(\mathrm{ED})_{it} \\ &\quad + \beta_{2i} \ln \mathrm{OF}\,(\mathrm{ED})_{it} + \varepsilon_{it}, \\ \ln \mathrm{VDS(ED)}_{it} &= \ln \beta_{0i} + \beta_{1i} \ln \mathrm{Zan}\,(\mathrm{ED})_{it} \\ &\quad + \beta_{2i} \ln \mathrm{OF}\,(\mathrm{ED})_{it} + \varepsilon_{it}, \end{aligned} \quad (17)$$

where $\mathrm{VDS}\,(\mathrm{ED})_{it}$ is the gross value added (in million UAH) for $i$th kind of economic activity in the $t$th time period per capita":

$\mathrm{Zan(ED)}_{it}$ is the number of population (thousand people) who are engaged in $i$th kind of economic activity in the $t$th time period;

$\mathrm{OF(ED)}_{it}$ is the value of fixed assets (in million UAH) for $i$th kind of economic activity in the $t$th time period; $\varepsilon_{it}$ is the random component;

$\beta_{0i}$, $\beta_{1i}$, $\beta_{2i}$, and $\rho_i$ are unknown parameters, which need to be quantitatively estimated.

The original data for model building data for 15 kinds of economic activity was used: agriculture; forestry and interlinked services; fishery (E1_E2); extractive industry;

reprocessing industry; energy, gas, and water production and distribution (E3_E4_E5); construction (E6); trade, car maintenance, and household goods; hotel and restaurant activities (E7_E8); transport and communication activities (E9); financial activity (E10); real estate operations, leasing, and engineering (E11); state or government management (E12); education (E13); health care and provision of social aid (E14); provision of communal and individual services (E15).

In addition, an analysis of the regional differences of effectiveness was performed for the investments into a type of economic activity for a determination of regional priorities for a distribution of the investment transfers. The model variants of panel data were considered and are defined as

$$\ln \mathrm{VDS}(\mathrm{ED})_{it}^{j} = \ln \beta_{0i}^{j} + \rho_i^j t + \beta_{1i}^{j} \ln \mathrm{Zan}(\mathrm{ED})_{it}^{j} + \beta_{2i}^{j} \ln \mathrm{IOK}(\mathrm{ED})_{it}^{j} + \varepsilon_{it}^{j} \quad (18)$$

$$\ln \mathrm{VDS}(\mathrm{ED})_{it}^{j} = \ln \beta_{0i}^{j} + \beta_{1i}^{j} \ln \mathrm{Zan}(\mathrm{ED})_{it}^{j} + \beta_{2i}^{j} \ln \mathrm{IOK}(\mathrm{ED})_{it}^{j} + \varepsilon_{it}^{j}, \quad (19)$$

where $\mathrm{VDS}(\mathrm{ED})_{it}^{j}$ is the gross value added per capita (in million UAH) for $j$th kind of economic activity for the $i$th region in the $t$th time period:

$\mathrm{Zan}(\mathrm{ED})_{it}^{j}$ is the number of population (thousand people) who are engaged in $j$th kind of economic activity for the $i$th region in the $t$th time period;

$\mathrm{IOK}(\mathrm{ED})_{it}^{j}$ is the rate of investment in fixed capital per capita (UAH) for $j$th kind of economic activity for the $i$th region in the $t$th time period;

$\varepsilon_{it}$ is the random component; $\beta_{0i}^{j}$, $\beta_{1i}^{j}$, $\beta_{2i}^{j}$, and $\rho_i^j$ are unknown parameters, which need to be quantitatively estimated.

The received industry functions will provide the possibility for an estimation of the technology effectively in some production investment and a potential growth in the gross added value.

From some alternative scenarios of financial regional policy, the compensating scenario (scenario 3) was considered, which provides a stimulation of economic growth for "problem" and stagnation regions and also for the region-"leaders," which have a slowdown of economic growth rates. At the same time, a transformation was being considered as the possibility for using of the distribution mechanisms from 2013. The original data for a scenario-form was being also considered for the forecast of some tax revenues, which were received based on a model of the imbalanced alignment of socio economic systems by use of tax levies. A pessimistic developed scenario also conducted this analysis, since this scenario allows an estimation of possibility for the formation of a "compensating" effect based upon a change of budget policy parameters.

In the alternative anticrisis scenario (scenario 4) a systematic financial support for the regions-recipients and regions-donors was considered. In the simulation some

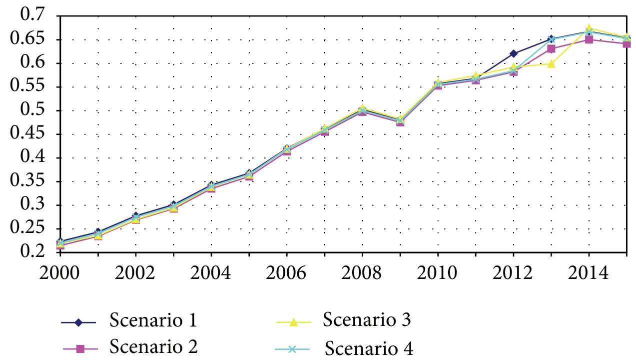



FIGURE 8: The integral indicator values of socioeconomic regional development.

investment transfers values in 2013 were accounted for, which were accepted as distribution parameters for the regional-development fund in the budget codecs and are oriented to the priority financial support for "problematic" territories. Correction of the distribution parameters of the investment transfers was carried out in 2014 with the purpose of warning of a cyclic recession in a dynamic of macroeconomic indicators (this recession is forecast in 2015), which is directed to financial support not only for "problem" territories, but also for the region-"leaders," for which a significant recession of the economic growth rates is observed.

The values of the integral indicator of the socioeconomic regional development, which are characterized by a developmental tendency of national economy in general, at the different scenarios of a fiscal policy, are shown in Figure 8.

Figure 8 shows that the realization of scenario 3 is formed by a forecast stagnation phase in a dynamic of the macroeconomic indicators, which is confirmed by the effectiveness of the accepted stabilization policy, allowing a rolling over of a growth phase for 2013-2014. Change of the budget policy parameters in 2014 can provide the possibility of decreasing the crisis depth in comparison with a base pessimistic scenario of budget insufficiency. The integral dynamic indicator for realization of scenario 4 is matched with a dynamic of an integral indicator in realization of scenario 1, accounted for with an optimistic forecast for tax revenues in a budget.

## 6. Results of the Research

The research highlights the following conclusions: analysis of the predictive dynamics of socioeconomic development territories, in the case of the implementation of an optimistic scenario of tax revenue, demonstrates the effectiveness of the adopted stabilization policy. The research serves to prevent the establishment of a crisis in the dynamics of macroeconomic indicators and indexes of regional development, to sustain a development phase. In addition, this research allows forming the "compensatory" effect of cutting the capacity of financial regulation of the developmental territories based on varying the parameters of the regional financial policy, which should be addressed to support "problematic" areas during the pessimistic scenario development of indicators of a budget system. As well as the results of this research allow to use the "basic" policy of levelling socioeconomic development of the regions "leaders," which are a significant slowdown for the economic growth tempos. The analysis of the coefficients of the instability of socioeconomic development shows a convergence trend of the stages of economic development territories, with different scenarios of development and a reduction in intergroup socioeconomic differentiation.

## 7. Conclusion

The developed scenario models offer the possibility for estimating the consistency for tax, budget, and investment policy. It also allows an increase in the quality of the data-analytical base for management decision-making with regard to the financial stability policy for regions and for the state. However, one promising direction of distribution of this inquiry is the structural and parametric adaptation of the proposed complex of the models. At the same, this complex can be used for estimation of a potential of the interregional collaboration and interaction. In addition, some other directions of this study will be able to detect of some possibilities for minimization of monetary values, which are re-connected with a stimulation of the territories development.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References


[1] I. G. Lukyanenko, *System Simulation of the Indicators of Budgetary System of Ukraine: Principles and Instruments*, PH Kyiv-Mohylanska Academy, 2004.

[2] P. O. Gourinchas and H. Rey, *External Adjustment, Global Imbalances, Valuation Effects*, 2013.

[3] L. Anselin and S. J. Rey, *Perspectives on Spatial Data Analysis*, Springer, Berlin, Germany, 2010.

[4] P.-P. Combes, M. Lafourcade, J.-F. Thisse, and J.-C. Toutain, "The rise and fall of spatial inequalities in France: a long-run perspective," *Explorations in Economic History*, vol. 48, no. 2, pp. 243–271, 2011.

[5] N. N. Lychkina, "Simulation modeling of regions' social and economic development in decision support systems," in *Proceedings of the System Dynamics Society Conference*, Albuquerque, NM, USA, 2009.

[6] G. Weisbrod, "Models to predict the economic development impact of transportation projects: historical experience and new applications," *Annals of Regional Science*, vol. 42, no. 3, pp. 519–543, 2008.

[7] H. Briassoulis, "Integrated economic- environmental-policy modeling at the regional and multiregional level: methodological characteristics and issues," *Growth & Change*, vol. 17, no. 3, pp. 22–34, 1986.

[8] J. H. Matis, T. R. Kiffe, T. I. Matis, and D. E. Stevenson, "Nonlinear stochastic modeling of aphid population growth," *Mathematical Biosciences*, vol. 198, no. 2, pp. 148–168, 2005.

[9] W. Qiu, X. Liu, and H. Li, "High-order fuzzy time series model based on generalized fuzzy logical relationship," *Mathematical Problems in Engineering*, vol. 2013, Article ID 927394, 11 pages, 2013.

[10] Y. Wang, Y. Zhang, F. Zhang, and J. Yi, "Robust quadratic regression and its application to energy-growth consumption problem," *Mathematical Problems in Engineering*, vol. 2013, Article ID 210510, 10 pages, 2013.

[11] K. V. Ketova, *Development of methods for the study and optimization of the development strategy of regional economic system [Ph.D. thesis]*, Izhevsk State Technical University, Izhevsk, Russia, 2008.

[12] V. A. Vasil'ev and V. I. Suslov, "Edgeworth equilibrium in a model of interregional economic relations," *Journal of Applied and Industrial Mathematics*, vol. 5, no. 1, pp. 130–143, 2011.

[13] V. A. Vasil'ev and V. I. Suslov, "On the unblockable states of multiregional economic systems," *Journal of Applied and Industrial Mathematics*, vol. 4, no. 4, pp. 578–587, 2010.

[14] V. Geyts, M. Kizim, T. Klebanova, and O. Chernyak, *Modeling of the Economic Security: State, Region, Enterprise*, Monograph, Publishing House INZHEK, Kharkiv, Ukraine, 2004.

[15] O. Chernyak and Y. Chernyak, "Modern challenges in governmental regulation of labour force migration in Ukraine," *Ekonomica*, vol. 91, no. 1, pp. 93–104, 2012.

[16] T. Toni and M. P. H. Stumpf, "Simulation-based model selection for dynamical systems in systems and population biology," *Bioinformatics*, vol. 26, no. 1, pp. 104–110, 2010.

[17] T. Klebanova, *The Uneven and Cyclical Dynamics of Socio-Economic Development of Regions: Assessment, Analysis, Prediction*, Monograph, Publishing House INZHEK, Kharkiv, Ukraine, 2012.

[18] T. Klebanova, L. Guryanova, Y. Daradkeh, and S. Kavun, "Approach to the assessment irregularity and cyclic dynamics of territorial development," *Asian Economic and Financial Review*, vol. 3, no. 12, pp. 1620–1641, 2013.

[19] T. Klebanova, S. Kavun, and L. Guryanova, "Models of assessment of inequality and skewness of social-economic systems development," *International Journal Biomedical Soft Computing and Human Sciences, Special Issue on Variational Bilevel Programming, Optimization Methods, and Applications to Economics*, vol. 18, no. 1, pp. 49–55, 2012.

[20] T. Klebanova and L. Chagovets, "Informational model for assessment of the most important indicators of social- economic regional non-uniform development," *The Advanced Science Journal*, no. 2, pp. 81–84, 2010.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 654293, 9 pages
http://dx.doi.org/10.1155/2014/654293

*Research Article*

# Selecting Large Portfolios of Social Projects in Public Organizations

**Igor Litvinchev,[1] Fernando Lopez-Irarragorri,[1] Nancy Maribel Arratia-Martínez,[1] and José Antonio Marmolejo[2]**

[1] *Faculty of Mechanical and Electrical Engineering, Nuevo Leon State University, 66451 Monterrey, NL, Mexico*
[2] *Faculty of Engineering, Anahuac University, 52786 Mexico City, MEX, Mexico*

Correspondence should be addressed to Fernando Lopez-Irarragorri; flopez65@gmail.com

Received 23 December 2013; Revised 11 March 2014; Accepted 11 March 2014; Published 23 April 2014

Academic Editor: Vyacheslav Kalashnikov



We address the portfolio selection of social projects in public organizations considering interdependencies (synergies) affecting project funds requirements and tasks. A mixed integer linear programming model is proposed incorporating the most relevant aspects of the problem found in the literature. The model supports both complete (all or nothing) and partial (a certain amount from a given interval of funding) resource allocation policies. Numerical results for large-scale problem instances are presented.

## 1. Introduction

One of the main management problems in many organizations and institutions in the public sector is to decide how to invest and manage funds available to support potential research and development projects in various areas [1]. A classification of the portfolio problem [2] distinguishes two types: (a) static and (b) dynamic. In (a), only projects proposals for funding are involved, while, in (b), at certain moments, some active projects are withdrawn from the portfolio and some inactive projects are activated.

Selecting portfolio of social projects is a periodical activity, involving a group of projects with social impact competing for financial support. This problem is categorized under the static type of portfolio problems.

It is important to highlight the differences between portfolio problems in public and private sectors. In contrast to public sector, the problem that occurs in private organizations is typically a dynamic one. Along with other differences, such as selection criteria and characteristics of the projects, in private sector, the projects can be evaluated more than once during the implementation. Moreover, criteria used to measure the portfolio in private sector are typically related to economic factors, such as expected sales, return of inversion or profit, and market situation. In the selection of portfolios of social projects, an important criterion is the social impact. Typically return of inversion or profit is not expected.

We refer to a social project as a group of tasks or activities consuming funds, carried out during a given period of time in one or more regions, with an impact on the objectives set by an organization, and focused on providing solution to problems or needs of the society.

Structural characteristics of the problem are as follows.

(1) There is a set of projects proposals competing only for funding. Every proposal is evaluated by a group of experts (reviewers) assigning to the proposal a value of its social impact (benefit).

(2) Every proposal has its own estimation of the funds required.

(3) The reviewers suggest a range of funding for each project.

(4) Each project belongs to a certain area of interest defined in the call for proposals. In particular, the area can be associated with geographical regions.

(5) The available budget is lower than the total amount of funds requested by all the proposals.

(6) Maximal and minimal value of funds that can be assigned to a given area are defined according to the organization rules.

(7) The total amount of funds requested by each proposal has to be divided between the tasks included in the proposal. Since it is difficult to set up a priori an exact value of funding for each task, a range of funding is defined, indicating the minimal and the maximal amounts of funds to be allocated to each task.

(8) The decision of funding is taken only once in a given period of time.

(9) All the projects are scheduled to start and finish in the same time period.

(10) All the tasks included in the project must start and finish in the same time period.

(11) There are interdependences between projects or tasks, resulting in synergies affecting the benefice and/or providing increase-decrease of required funds.

(12) According to organization rules, there may be certain relationships between projects/tasks, which restrict a number of projects/tasks included in the portfolio.

Below we summarize principal characteristics of our approach to portfolio selection problem.

*(i) Characterization of Projects through a Set of Tasks.* It is natural to see the project as a set of tasks requiring funds for their implementation. If so, each of these tasks must make a contribution to the impact of the project, contributing to the quality measure of the portfolio. Thus it is necessary to include the amount of funds requested by the task, as well as an index of its importance for achieving objectives of the project.

*(ii) Flexibility of the Funds Allocated to Tasks.* In most of references on portfolio selection, the projects are either supported by the complete amount requested in the proposal or not supported at all. However, in many real-word problems, a part of the requested amount can also be assigned. Taking this into account, we assume that reviewers (if necessary, supported by experts) are able to analyze if the request is overestimated (underestimated). This way, an interval of possible funds assignment is defined allowing flexible assignment within the interval. Similarly, for the tasks. Hence, the funding of the project can vary according to funds assigned to its tasks.

*(iii) Interdependences between Projects or Project Tasks.* Another important factor to consider is the existence of synergies between tasks produced by interdependencies in tasks and/or projects. According to Rungi [3], projects synergies appear when the measure of the quality of all the projects is different from the sum of the measures of individual projects. In this paper, we consider the interdependencies between tasks of two or more projects producing synergies of the following types: (1) synergies of benefit or impact, (2) synergies which result in an increase in funds, (3) synergies which results in a decrease in funds, and (4) technical synergies. These are defined for groups of synergies indicating that only a subset of them can be active. Technical synergies are used to have greater control over the synergies of any kind.

The following assumptions are used in this paper.

(i) There is a person (or persons) referred to as the decision maker (DM) and representing preferences and priorities of the organization. Typically the DM is represented by reviewers of the project proposals.

(ii) The DM is able to express the advantages/disadvantages providing by synergies.

(iii) Each project can be divided into a number of tasks requiring funds.

(iv) The tasks of each project have different impacts on the objectives of the project they belong to. The value of impact is provided in the proposal and is evaluated by reviewers.

(v) The reviewers are able to define the interval of funds allocated for each task and project.

Bearing in mind the characteristics stated above, the solution of the problem consists in making the following decisions:

(1) selection of projects and corresponding tasks to be included in the portfolio,

(2) allocation of funds to tasks of selected projects,

(3) evaluation of synergetic effects produced by interdependences between projects/tasks included in the portfolio.

## 2. Some Background

Tables 1 and 2 summarize some characteristics of portfolio problems considered in the references.

Table 1 shows that there is a tendency to consider more than one objective in the evaluation of the portfolio.

These objectives can be of general form as in [4, 5] or may represent specific measures associated with a set of attributes. Portfolio problems in the private sector are focused mainly on inversion related objectives and sales, among other market related issues. In social or public areas, usually the aim is to maximize the quality of the portfolio related to the social impact of the projects selected.

The total allocation of resources means that once a project was selected for inclusion in the portfolio, it receives all requested resources. However, in many real situations, it is important to consider partial allocations of resources when a project receives less than what was requested.

In the work of Wang and Hwang [6], they consider the uncertain character of the project cost and use fuzzy sets concept for modeling. Though resource assignment was not considered explicitly in their model, a partial assignment is implied.

Table 2 presents some characteristics of social portfolios. Balance constraints may be established by organizations to set upper/lower bounds for funds designated, in particular, for R&D areas. This way, the portfolio can be balanced to assure at least minimal funding for some specific areas.

Table 1: General characteristics.

| Paper | Objectives | Maximum number of projects in portfolio | Resource allocation policy | Solution methodology |
|---|---|---|---|---|
| Ghasemzadeh and Archer, 2000 [18] | General benefits integrated in a weighted sum | 10 | Complete | Integer linear programming model adjusted interactively |
| Stummer and Heidenberger, 2003 [4] | General benefit and resources consumption by period of scheduling | 30 | Complete | Integer linear programming model to get efficient solutions; the best compromise solution is selected interactively |
| Wang and Hwang, 2007 [6] | Total benefit obtained by summing expected returns of the projects minus total inversion | 20 | Partial | Fuzzy model |
| Gutjahr et al., 2008 [7] | Economic gains and strategic gains | 18 | Partial | Nonlinear mixed integer programming model, greedy heuristic, and two alternative metaheuristics (ACO and GA) |
| Carazo et al., 2010 [5] | General benefit categories | 90 | Complete | Metaheuristics (SS-PPS, Scatter Search) |
| Gutjahr et al., 2010 [8] | Economic benefits and competence benefits | 18 | Partial | Mixed integer linear model, two metaheuristics (NSGA-II and P-ACO) |
| Litvinchev et al., 2010 [15] | Portfolio quality and number of projects in portfolio integrated in a weighted sum | 25000 | Partial | Mixed integer linear programming model; a compromise solution is obtained interactively |
| Litvinchev et al., 2011 [16] | Portfolio quality and number of projects in portfolio integrated in a weighted sum | 10000 | Partial | Mixed integer linear programming model; a compromise solution is obtained interactively |
| Gutjahr and Froeschl, 2013 [9] | Expected return minus outsourced costs | 15 | Partial | Metaheuristic method (S-VNS) and the Frank-Wolfe algorithm |

Table 2: Specific characteristics.

| Paper | Balance constraints | Tasks | Interdependencies | Scheduling | Risk |
|---|---|---|---|---|---|
| Ghasemzadeh and Archer, 2000 [18] | No | No | General | Multiple periods | As an attribute of portfolio |
| Stummer and Heidenberger, 2003 [4] | No | No | Benefit, increment, or decrement in resources consumption | Multiple periods | No |
| Wang and Hwang, 2007 [6] | No | No | No | Single period | Risk aversion |
| Gutjahr et al., 2008 [7] | No | Yes | No | Multiple periods | No |
| Carazo et al., 2010 [5] | No | No | Benefit, increment, or decrement in resources consumption | Multiple periods | As an attribute in portfolio |
| Gutjahr et al., 2010 [8] | No | Yes | No | Multiple periods | No |
| Litvinchev et al., 2010 [15] | No | No | No | Single period | Neutral position |
| Litvinchev et al., 2011 [16] | Yes | Yes | Benefit, increment, or decrement in resources consumption | Single period | Neutral position |
| Gutjahr and Froeschl, 2013 [9] | No | Yes | Yes | Multiple periods | Risk aversion by penalizing the variance of the return |

Another important characteristic of the social portfolio problems is considering a project as a set of tasks, each with a certain impact on the project objectives. In [7–9], the projects are represented by tasks that require human resources and that have relations of precedence in the planning horizon. In [10, 11], tasks are also used in project portfolio scheduling. It was pointed out in [12] that, by representing projects by tasks, it is possible to have a control of expenses or resources consumption or even of the impact of projects on portfolio impact measures.

Since, in social area, usually only one time period is considered, only a single evaluation of proposals is implemented and a single decision on funding is taken. In contrast, in private area, proposals can be reevaluated to continue funding or even stop funding and choose another project for support if a project is no longer attractive for investment.

## 3. The Model

It is assumed that there is a set of projects competing for funding in a number of areas of research and development, such that every project belongs to a single area. The project consists of a number of tasks and the project is supported if at least one of its tasks is supported.

The project is funded sufficiently if it receives support in a certain interval of funding. The corresponding benefit function increases linearly on this interval and is zero for funding below the minimal value (see [13] for details). Sufficient funding for tasks is defined similarly.

To represent interdependency between different projects/tasks, the concept of synergy is used. Generally speaking, synergy arises when the total indicator corresponding to elements (projects, tasks, etc.) from the synergy is different from the sum of indicators corresponding to the same elements considered independently. More specifically, synergy is interpreted as a set of tasks corresponding to certain projects that, being supported sufficiently, produce an effect of changing the benefit or increase/decrease of funds. In [4], it was assumed that synergy is activated if all its elements are funded sufficiently; that is, cardinality of synergy coincides with that of the set of the elements funded. In this paper, we use a more flexible definition of synergy. Similar to [5, 14], we say that synergy is activated if a number of its elements funded sufficiently lie within certain bounds.

Three types of synergies are considered. Benefit synergies, being activated, result in increase/decrease of the overall benefit of the portfolio. Resource synergies result in increase/decrease of the overall funding to elements of synergy. It is assumed that resource synergies are defined for tasks corresponding to projects of the same area. Technical synergies aimed to limit the number of activated synergies of certain type. To state the model mathematically, the following notation is used.

### 3.1. Sets and Parameters

$J$: set of projects competing for financial support, $j = 1, 2, 3, \dots, |J|$.

$K$: set of areas of research and development, its indices $k = 1, 2, 3, \dots, |K|$.

$J_k$: set of projects belonging to area $k$.

$I$: set of tasks $i = 1, 2, 3, \dots |I|$.

$C$: set of synergies, $s = 1, 2, 3, \dots |C|$.

$C^s$: set of elements (pairs) of synergy $s$, $C^s = \{(j_1, i_1), (j_2, i_2), \dots, (j_{|C^s|}, i_{|C^s|})\}$.

$B$: set of benefit type synergies, $B \subset C$.

$L$: set of synergies of funds reduction, $L \subset C$.

$H$: set of synergies of funds enlargement, $H \subset C$.

$\tau$: set of technical synergies.

$\eta^s$: value of synergetic effect of funds reduction by synergy $s$.

$\lambda^s$: value of synergetic effect of funds enlargement by synergy $s$.

$\nu^s$: value of synergetic effect of benefit by synergy $s$.

$w_j$: the social impact of the project $j$.

$\rho_{ji}$: relative importance of the task $i$ of the project $j$.

$PG$: available budget.

$R^-_{ji}$, $R^+_{ji}$: minimum and maximum amount to fund the task $j, i$.

$P^-_k$, $P^+_k$: minimum and maximum amount of funds to the area $k$.

$M^-_j$, $M^+_j$: minimum and maximum amount to fund the project $j$.

$m^{s-}$, $m^{s+}$: minimum and maximum number of tasks to enable synergy $s$.

$E^-_T$, $E^+_T$: minimum and maximum number of synergies to enable technical synergy $T$.

$d^-_j$, $d^+_j$: minimum and maximum number of projects to fund by area.

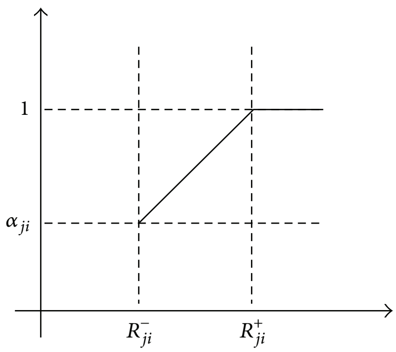



Figure 1: Representation of the piecewise linear function.

*3.2. Variables.* Let $x_{ji}$ be an amount of funds assigned to task $i$ of project $j$. Consider

$$y_j = \begin{cases} 1, & \text{the project } j \text{ is enough supported} \\ 0, & \text{otherwise} \end{cases}$$

$$z_{ji} = \begin{cases} 1, & \text{the task } j, i \text{ is enough supported} \\ 0, & \text{otherwise} \end{cases}$$

$$\sigma_1^s = \begin{cases} 1, & \text{synergetic set has cardinality} \geq m^{s-} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

$$\sigma_2^s = \begin{cases} 1, & \text{synergetic set has cardinality} \leq m^{s+} \\ 0, & \text{otherwise} \end{cases}$$

$$\sigma^s = \begin{cases} 1, & \text{synergy } s \text{ is activated}, \sigma^s = \sigma_1^s + \sigma_2^s - 1 \\ 0, & \text{otherwise.} \end{cases}$$

*3.3. Objectives.* Two objectives are considered: (1) represents the quality of the portfolio in a similar way as in Litvinchev et al. [15] and (2) represents the total amount of funded projects:

$$\sum_{j \in J} w_j \left( \sum_{i \in I} \left( a_{ji} z_{ji} + b_{ji} x_{ji} \right) \cdot \rho_{ji} \right) + \sum_{s \in B} \nu^s \sigma^s \tag{2}$$

$$\sum_{j \in J} y_j. \tag{3}$$

We consider the number of funded projects as an objective since if we have two portfolios with equal quality and resource consumption, the portfolio with more projects is preferable.

In (2), the parameters $a_{ji}$ and $b_{ji}$ are defined similarly to Litvinchev et al. [15] but adjusted to the representation of projects by tasks. Essentially, the predicate "a task is sufficiently funded" is modeled with a degree of truth represented by a piecewise linear function $\mu_{ji}$, Figure 1 increasing on $[R_{ji}^-, R_{ji}^+]$, such that $\mu_{ji}(R_{ji}^+) = \alpha_{ji}$ and $\mu_{ji}(R_{ji}^-) = 1$ [15].

In this way, the parameters $a_{ji}$ and $b_{ji}$ are defined such that $a_{ji} + b_{ji} x_{ji} = 1$ for $x_{ji} = R_{ji}^+$ and $a_{ji} + b_{ji} x = \alpha_{ji}$ for $x_{ji} = R_{ji}^-$.

Thus,

$$a_{ji} = \alpha_{ji} - \frac{R_{ji}^- \left(1 - \alpha_{ji}\right)}{R_{ji}^+ - R_{ji}^-}, \qquad b_{ji} = \frac{\left(1 - \alpha_{ji}\right)}{R_{ji}^+ - R_{ji}^-}. \tag{4}$$

For the case $R_{ji}^- < R_{ji}^+$, the expressions in (4), as well as the coefficients of the objective (2), are well defined. If $R_{ji}^- = R_{ji}^+$ or are very close to each other, the objective (2) has to be modified. For $R_{ji}^- = R_{ji}^+$, we may set $a_{ji} = 1$, $b_{ji} = 0$, such that the corresponding benefit is either 0 or $a_{ji}$. Note that, in this case by constraint (8) below, the task is either supported completely ($z_{ji} = 1$) or not supported at all ($z_{ji} = 0$).

*3.4. Constraints.* The constraints of the model are

$$\sum_{j \in J} \sum_{i \in I} x_{ji} \leq P_G \tag{5}$$

$$P_k^- \leq \sum_{j \in J_k} \sum_{i \in I} x_{ji} \leq P_k^+, \quad k \in K \tag{6}$$

$$M_j^- y_j \leq \sum_{i \in I} x_{ji} \leq M_j^+ y_j, \quad j \in J \tag{7}$$

$$R_{ji}^- z_{ji} \leq x_{ji} \leq R_{ji}^+ z_{ji}, \quad \left\{ \begin{array}{c} i \in I, j \in J \\ (j,i) \notin L \cup H \end{array} \right. \tag{8}$$

$$y_j \leq \sum_{i \in I} z_{ji}, \quad j \in J \tag{9}$$

$$\sum_{i \in I} z_{ji} \leq |I| \, y_j, \quad j \in J \tag{10}$$

$$\sum_{(j,i) \in C^s} z_{ji} - m^{s-} + 1 \leq \left|C^s\right| \sigma_1^s, \quad s \in C \tag{11}$$

$$\left|C^s\right| \sigma_1^s \leq \sum_{(j,i) \in C^s} z_{ji} - m^{s-} + \left|C^s\right|, \quad s \in C \tag{12}$$

$$m^{s+} - \sum_{(j,i) \in C^s} z_{ji} + 1 \leq \left|C^s\right| \sigma_2^s, \quad s \in C \tag{13}$$

$$\left|C^s\right| \sigma_2^s \leq m^{s+} - \sum_{(j,i) \in C^s} z_{ji} + \left|C^s\right|, \quad s \in C \tag{14}$$

$$\sigma^s = \sigma_1^s + \sigma_2^s - 1, \quad s \in C \tag{15}$$

$$\sum_{(j,i) \in C^s} x_{ji} \leq \sum_{(j,i) \in C^s} R_{ji}^+ z_{ji} - \eta^s \sigma^s, \quad s \in L \tag{16}$$

$$\sum_{(j,i) \in C^s} x_{ji} \geq \sum_{(j,i) \in C^s} R_{ji}^- z_{ji} + \lambda^s \sigma^s, \quad s \in H \tag{17}$$

$$\sum_{(j,i) \in C^s} x_{ji} \geq \sum_{(j,i) \in C^s} R_{ji}^- z_{ji} - \eta^s \sigma^s, \quad s \in L \tag{18}$$

$$\sum_{(j,i) \in C^s} x_{ji} \leq \sum_{(j,i) \in C^s} R_{ji}^+ z_{ji} + \lambda^s \sigma^s, \quad s \in H \tag{19}$$

$$E_T^- \leq \sum_{s \in T} \sigma^s \leq E_T^+, \quad T \in \tau \tag{20}$$

$$x_{ji} \geq 0, \quad y_j, z_{ji}, \sigma^s, \sigma_1^s, \sigma_2^s \in \{0, 1\}. \tag{21}$$

Constraints (5)–(8) are typical budged constraints, defined at portfolio, area, project, and task levels, respectively. Constraints (9)-(10) relate tasks to projects. All these constraints are defined in Litvinchev et al. [16].

Constraints (11)–(15) represent the activation of synergies. Defining synergies at the task level provides more flexibility comparing with the definition of synergies at the project level. Here a synergy is activated if a subset of tasks in $C^s$ is funded with cardinality less than a number $m^{s+}$ and greater than a $m^{s-}$ (see also [5]).

If $m^{s+} = m^{s-}$, this implies that the synergy $s$ is activated only if all the tasks associated with that synergy are supported. In (11)–(14), variables $\sigma_1^s$ and $\sigma_2^s$ correspond to lower and upper limits, respectively. We define $\sigma^s$ in terms of the previous two variables to verify that the synergy $s$ is activated. Variables $\sigma_1^s$ and $\sigma_2^s$ cannot be equal to zero simultaneously, since that would imply that the sum $\sum_{(j,i)\in C^s} z_{ji}$ is less than $m^{s-}$ and greater than $m^{s+}$. This is impossible since $m^{s-} \leq m^{s+}$. Moreover, our formulation for the restriction (20) eliminates the nonlinearity of the model of Carazo et al. [5]. Constraints (16)–(19) represent resources synergies.

Technical synergies are defined to keep control over the synergies of tasks that could be activated. The aim is to limit a number of active synergies from a group of them. We call $T$ the set containing the indices of synergies to be limited; that is, $T = \{s_1, s_2, s_3, \ldots, s_{|T|}\}$, so that only a number less than or equal to $E_T^+$ and/or greater than or equal to $E_T^-$ of them can be active. The set of all $T$'s is $\tau$. In our model, technical constraints are represented by constraint (20). This representation allows incorporating more than two groups of mutually exclusive groups of tasks.

*Other Constraints*. In some cases, it is desirable to balance the portfolio for the number of projects supported by area. That is,

$$d_j^- \leq \sum_{j\in J_k} y_j \leq d_j^+, \tag{22}$$

where $d_j^-$ and $d_j^+$ are respective lower and upper bounds.

## 4. Numerical Experiments

Different groups of instances were generated varying the number of projects proposals, tasks, and synergies. The objective of the first part of the numerical study was to test the tractability of the model for large-scale instances. The commercial software CPLEX 12.5 was used for optimization running on a server DELL PowerEdge 2950 with 8 cores. Six classes of instances were generated each having 15 replications: (1) 100 projects, 2 areas, 5 tasks, and 5 synergies; (2) 500 projects, 2 areas, 15 tasks, and 40 synergies; (3) 1000 projects, 2 areas, 15 tasks, and 80 synergies; (4) 1000 projects, 2 areas, 15 tasks, and 200 synergies; (5) 10000 projects, 4 areas, 15 tasks, and 0 synergies; and (6) 25000 projects, 4 areas, 1 task, and 0 synergies. The instances were coded according to Projects Areas Tasks Synergies, such that P100A2T5S5 corresponds to the first class.

Table 3: CPU time.

| Instance | Mean CPU time (sec.) | Std. dev. |
|---|---|---|
| P100A2T5S5 | 0.208 | 0.034 |
| P500A2T15S40 | 6.397 | 1.035 |
| P1000A2T15S80 | 59.158 | 17.462 |
| P1000A2T15S200 | 328.233 | 132.692 |
| P10000A4T15S0 | 588.47 | 223.18 |
| P25000A4T1S0 | 8.67 | 2.11 |

Table 3 shows the results for one run of the model using a weighted sum of the objectives (weighting parameters of 0.5 for both objectives were used in all experiments). For instances with less than 1000 projects, the running time was on average less than 1 min. For large-scale instances with more than 1000 projects, the running time increased in the worst case up to 10 min.

From Table 3, we may conclude that the number of projects is not a critical factor for the increasing of CPU time. However the number of tasks and, especially, the number of synergies represent a key factor for the complexity of the problem, as can be observed in instances P25000A4T15S0 and P1000A4T15S200. Bearing in mind that the model is used for long-term decisions, we may conclude that the model provides decisions in a reasonable time. For more numerical experiments for very large-scale instances, see Irarragori and Martínez [17].

The second part of the numerical experiment was conducted to see the behavior of the model subject to changes in some parameters. In particular, upper and lower bounds for the funds to be assigned for projects and tasks were of special interest. Three classes of 15 instances were considered (P100A2T5S5, P500A2T15S40, and P1000A2T15S80).

First, all instances were solved and the respective values of the objectives (quality of the portfolio and number of projects) and the solution time were stored. Then we reduced the interval $[R_{ji}^-, R_{ji}^+]$ of possible funding for each task of the project. This was done by shifting the minimal amount $R_{ji}^-$ towards the maximal value $R_{ji}^+$ keeping the later constant. The new interval was obtained in the form $[R_{ji}^- + \gamma(R_{ji}^+ - R_{ji}^-), R_{ji}^+]$, where $\gamma \in [0, 1]$ represents the index of reduction. In the experiment 7 levels of reduction were considered corresponding to different values of $\gamma$ (in %): 10%, 30%, 50%, 90%, 95%, and 100%. Reduction 100% ($\gamma = 1$) corresponds to $R_{ji}^- = R_{ji}^+$. In this case by constraint (8), the task is either supported completely ($z_{ji} = 1$) or not supported at all ($z_{ji} = 0$). The reduction of funds was applied to a different number of projects. We selected randomly 25%, 50%, 75%, and 100% of all the projects being subject to funds reduction.

The results of the corresponding computational experiment are summarized in Tables 4–6. The indicators presented in these tables are the average (over all problem instances) reductions in number of projects supported and in the quality measure. More specifically, the quality reduction indicator (QR) in Tables 4(a), 5(a), and 6(a) was defined as QR = $(1 - NQ/Q) \cdot 100\%$, where Q denotes the quality measure without

Table 4: QR (a) and PR (b) for instances of P100A2T5S5.

(a)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 0.847 | 1.138 | 1.753 | 5.913 | 5.655 | 5.816 |
| 50 | 0.907 | 2.367 | 3.401 | 9.925 | 9.757 | 9.851 |
| 75 | 1.858 | 4.217 | 5.731 | 13.018 | 12.741 | 12.890 |
| 100 | 1.652 | 4.659 | 7.650 | 15.259 | 16.043 | 16.784 |

(b)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 0.816 | 2.961 | 5.621 | 11.173 | 11.664 | 11.749 |
| 50 | 1.968 | 6.348 | 11.636 | 19.274 | 20.020 | 20.689 |
| 75 | 2.637 | 9.438 | 17.877 | 29.812 | 30.555 | 30.965 |
| 100 | 4.635 | 13.324 | 23.844 | 39.151 | 39.900 | 40.994 |

Table 5: QR (a) and PR (b) for instances of P500A2T15S40.

(a)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 0.169 | 0.680 | 0.346 | 8.360 | 9.217 | 9.966 |
| 50 | 0.869 | 0.902 | 1.488 | 12.197 | 12.930 | 13.535 |
| 75 | 1.100 | 1.787 | 3.383 | 15.283 | 16.404 | 16.971 |
| 100 | 1.344 | 2.729 | 5.038 | 18.564 | 20.431 | 20.726 |

(b)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 1.512 | 4.066 | 7.152 | 12.771 | 12.711 | 13.348 |
| 50 | 2.401 | 8.149 | 13.050 | 21.857 | 23.123 | 23.555 |
| 75 | 3.726 | 11.790 | 19.031 | 31.944 | 33.467 | 34.114 |
| 100 | 5.279 | 15.393 | 24.978 | 41.012 | 43.177 | 43.922 |

Table 6: QR (a) and PR (b) for instances of P1000A2T15S80.

(a)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 0.770 | 1.066 | 1.614 | 8.833 | 9.123 | 9.236 |
| 50 | 0.903 | 2.127 | 3.355 | 12.525 | 13.251 | 13.664 |
| 75 | 1.178 | 3.240 | 4.806 | 16.196 | 17.662 | 17.930 |
| 100 | 1.594 | 4.360 | 6.217 | 19.996 | 21.760 | 22.157 |

(b)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 0.811 | 3.744 | 6.172 | 12.461 | 12.882 | 13.047 |
| 50 | 2.386 | 7.456 | 12.476 | 22.131 | 23.036 | 23.546 |
| 75 | 3.614 | 11.195 | 18.733 | 31.243 | 32.798 | 33.440 |
| 100 | 4.849 | 14.914 | 24.714 | 40.698 | 42.640 | 43.480 |

Table 7: Relative change of CPU time for P100A2T5S5 (a), P500A2T15S40 (b), and P1000A2T15S80 (c).

(a)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 5.77 | 10.10 | 7.69 | 25.00 | 20.67 | 20.67 |
| 50 | −1.44 | 13.46 | 16.83 | 22.12 | 30.77 | 29.33 |
| 75 | 2.88 | 0.96 | 34.62 | 11.54 | 13.46 | 21.63 |
| 100 | 5.29 | 15.87 | 42.31 | 53.37 | 26.92 | 8.17 |

(b)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 0.86 | 41.30 | 70.60 | 88.93 | 40.19 | 39.91 |
| 50 | 11.46 | 63.87 | 112.47 | 108.32 | 57.56 | 30.31 |
| 75 | 34.23 | 82.15 | 166.30 | 149.74 | 101.00 | 70.17 |
| 100 | 45.24 | 153.93 | 236.38 | 349.07 | 221.23 | 182.21 |

(c)

| Projects with reduction (in %) | Reduction index $\gamma$ (in %) | | | | | |
|---|---|---|---|---|---|---|
| | 10 | 30 | 50 | 90 | 95 | 100 |
| 25 | 0.85 | −23.97 | −19.04 | −18.20 | −13.98 | −12.55 |
| 50 | 5.24 | −38.80 | 8.84 | −17.29 | −36.57 | −28.56 |
| 75 | 6.19 | −11.91 | 86.36 | 0.44 | −30.70 | −45.96 |
| 100 | 16.25 | 65.65 | 165.24 | 44.41 | 63.64 | 56.82 |

funds reduction, while $NQ$ states the quality measure after funds reduction. Similarly, the project reduction indicator (PR) in Tables 4(b), 5(b), and 6(b) was defined as $\mathrm{PR} = (1 - NP/P) \cdot 100\%$, where $P$ and $NP$ denote the number of projects before and after funds reduction, respectively.

Observing the columns in Tables 4–6, we may conclude that both QR and PR increase monotonously as long as the number of projects with funds reduction is increased. This holds for all problem instances and all indices $\gamma$ of funds reduction. That is, applying funds reduction to more projects results in stronger decrease for both objectives of the problem, the quality, and the number of projects in the portfolio.

Considering the rows in Tables 4–6, we see that, for the fixed number of projects with funds reduction, both QR and PR typically increase as long as the value of the funds reduction index $\gamma$ is increased. However, this is not always the case. For example, comparing the first 3 rows in Table 4(a), we see that the QR slightly decreases when $\gamma$ changes from 90% to 95%. That is, reducing the funds results in increase of the quality. However, this is "compensated" by decrease in the number of projects supported.

Generally speaking, the weighted criterion (the weighted sum of the quality and the number of projects) decreases as long as the funds reduction $\gamma$ increases. This holds since solution to the weighted problem (2)–(22) obtained for a smaller funding interval remains feasible for the same problem with a larger funding interval. But this does not mean that both terms, quality and projects, have to be decreased. An increase of one term can be compensated by a larger decrease of the other.

Table 7 presents the values of indicator characterizing relative changes of CPU time due to funds reduction. This indicator was defined as $\mathrm{TR} = (NT/T - 1)100\%$, where $T$ is CPU time without funds reduction and $NT$ corresponds to CPU time after funds reduction. Note that a negative value of TR indicates that $NT$ is less than $T$. The average (over 15 problem instances) value of the indicator is presented.

As we can see from Table 7, if funds reduction is applied to all projects (the last row in tables), then reducing interval of funding always results in increase of CPU time (all values in the last row are positive). That is, instances with complete (all or nothing) type of funding for all projects are more difficult to solve for this model. However, if funds reduction is applied only to a part of projects, CPU time may either increase or decrease (negative values in the table). Comparing with the values of CPU time presented in Table 3, we may conclude that even if CPU time increases, it is still reasonably low.

## 5. Conclusions

This paper studies portfolio selection problem for R&D projects in public organizations. Two objectives are considered: benefit (the quality of portfolio related to the social impact of projects) and number of projects supported. It is assumed that there is a set of projects competing for funding in a number of areas of R&D, such that every project belongs to a single area. The project consists of a number of tasks and the project is supported if at least one task is supported. The project/task is funded sufficiently if it receives support in a certain interval of funding. The interdependency between projects/tasks is modeling using the general concept of synergy. A synergy is active if a number of its elements supported sufficiently are within certain bounds. Three types of synergies are considered, affecting the benefit, changing resource consumptions, and the so-called technical synergies, aimed at limiting the number of active synergies. A corresponding mixed integer linear model is presented providing solutions to large-scale instances in a reasonable time.

We demonstrated that, within the proposed model, the project/task can be supported, completely, receiving the funds requestedor, partially, receiving funds within a certain interval; that is, the concept of sufficient (continuous) funding was implemented.

The computational experiment demonstrates that increasing the number of projects subject to funds reduction results in decrease for both objectives. However, if the number of projects subject to funds reduction is fixed but the strength of funds reduction is increased, one of the objectives may increase at the expense of a larger decrease of another.

Only one type of the resources (funds) was considered in the model. An interesting direction for future research is studying problems with multiple resources, where synergetic effects can arise not only at the level of projects/tasks, but also at the level of resources.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was partially supported by Grants from RFBR, Russia (12_01_00893_a), and CONACYT, Mexico (167019). The authors would like to thank the anonymous referees for their constructive comments and suggestions.

## References

[1] J. N. Castillo, I. V. Lopez, E. F. Gonzalez, and E. L. Cervantes, "Aplicación de metaheuristicas multiobjetivo a la solución de problemas de cartera de proyectos públicos con una valoración multidimensional de su impacto," *Gestion y Politica Publica*, vol. 20, no. 2, pp. 381–432, 2011.

[2] H. Eilat, B. Golany, and A. Shtub, "Constructing and evaluating balanced portfolios of R&D projects with interactions: a DEA based methodology," *European Journal of Operational Research*, vol. 172, no. 3, pp. 1018–1039, 2006.

[3] M. Rungi, "Visual representation of interdependencies between projects," in *Proceedings of the 37th International Conference on Computers and Industrial Engineering*, pp. 1061–1072, Alexandria, Egypt, October 2007.

[4] C. Stummer and K. Heidenberger, "Interactive R&D portfolio analysis with project interdependencies and time profiles of multiple objectives," *IEEE Transactions on Engineering Management*, vol. 50, no. 2, pp. 175–183, 2003.

[5] A. F. Carazo, T. Gómez, J. Molina, A. G. Hernández-Díaz, F. M. Guerrero, and R. Caballero, "Solving a comprehensive model for multiobjective project portfolio selection," *Computers & Operations Research*, vol. 37, no. 4, pp. 630–639, 2010.

[6] J. Wang and W.-L. Hwang, "A fuzzy set approach for R&D portfolio selection using a real options valuation model," *Omega*, vol. 35, no. 3, pp. 247–257, 2007.

[7] W. J. Gutjahr, S. Katzensteiner, P. Reiter, C. Stummer, and M. Denk, "Competence-driven project portfolio selection, scheduling and staff assignment," *Central European Journal of Operations Research*, vol. 16, no. 3, pp. 281–306, 2008.

[8] W. J. Gutjahr, S. Katzensteiner, P. Reiter, C. Stummer, and M. Denk, "Multi-objective decision analysis for competence-oriented project portfolio selection," *European Journal of Operational Research*, vol. 205, no. 3, pp. 670–679, 2010.

[9] W. J. Gutjahr and K. A. Froeschl, "Project portfolio selection under uncertainty with outsourcing opportunities," *Flexible Services and Manufacturing Journal*, vol. 25, no. 1-2, pp. 255–281, 2013.

[10] C. M. D. M. Mota, A. T. de Almeida, and L. H. Alencar, "A multiple criteria decision model for assigning priorities to activities in project management," *International Journal of Project Management*, vol. 27, no. 2, pp. 175–181, 2009.

[11] U. Beşikci, Ü. Bilge, and G. Ulusoy, "Resource dedication problem in a multi-project environment," *Flexible Services and Manufacturing Journal*, vol. 25, no. 1-2, pp. 206–229, 2013.

[12] R. Zeynalzadeh and A. Ghajari, "A framework for project portfolio selection with risk reduction approach," *African Journal of Business Management*, vol. 5, no. 26, pp. 10474–10482, 2011.

[13] E. Fernandez, F. Lopez, J. Navarro, and I. Vega, "An integrated mathematical-computer approach for R&D project selection in large public organizations," *International Journal of Mathematics in Operational Research*, vol. 1, no. 3, pp. 372–396, 2009.

[14] A. F. Carazo, T. Gómez, and F. Pérez, "Análisis de los principales aspectos que afectan la decisión de selección y planificación de carteras e proyectos," *Revista Electrónica de Comunicaciones y Trabajos de ASEPUMA*, vol. 12, pp. 123–140, 2012.

[15] I. S. Litvinchev, F. López, A. Alvarez, and E. Fernández, "Large-scale public R&D portfolio selection by maximizing a biobjective impact measure," *IEEE Transactions on Systems, Man, and Cybernetics A: Systems and Humans*, vol. 40, no. 3, pp. 572–582, 2010.

[16] I. Litvinchev, F. López, H. J. Escalante, and M. Mata, "A milp biobjective model for static portfolio selection of R&D projects with synergies," *Journal of Computer and Systems Sciences International*, vol. 50, no. 6, pp. 942–952, 2011.

[17] F. L. Irarragori and N. M. A. Martínez, "R&D project portfolio selection in public organizations," Tech. Rep. PISIS-2013-01, Graduate Program in Systems Engineering, UANL, San Nicolás de los Garza, México, 2013.

[18] F. Ghasemzadeh and N. P. Archer, "Project portfolio selection through decision support," *Decision Support Systems*, vol. 29, no. 1, pp. 73–88, 2000.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 750232, 13 pages
http://dx.doi.org/10.1155/2014/750232

*Research Article*

# A Heuristic Procedure for a Ship Routing and Scheduling Problem with Variable Speed and Discretized Time Windows

**Krystel K. Castillo-Villar,[1] Rosa G. González-Ramírez,[2] Pablo Miranda González,[2] and Neale R. Smith[3]**

[1] *Department of Mechanical Engineering, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA*
[2] *Industrial Engineering School, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile*
[3] *Quality and Manufacturing Center, Tecnológico de Monterrey, Campus Monterrey, Mexico*

Correspondence should be addressed to Rosa G. González-Ramírez; rosa.gonzalez@ucv.cl

Received 22 November 2013; Revised 5 February 2014; Accepted 11 February 2014; Published 8 April 2014

Academic Editor: Sergii V. Kavun



This paper develops a heuristic algorithm for solving a routing and scheduling problem for tramp shipping with discretized time windows. The problem consists of determining the set of cargoes that should be served by each ship, the arrival, departure, and waiting times at each port, while minimizing total costs. The heuristic proposed is based on a variable neighborhood search, considering a number of neighborhood structures to find a solution to the problem. We present computational results, and, for comparison purposes, we consider instances that can be solved directly by CPLEX to test the performance of the proposed heuristic. The heuristics achieves good solution quality with reasonable computational times. Our computational results are encouraging and establish that our heuristic can be utilized to solve large real-size instances.

## 1. Introduction

In light of the phenomenon of globalization, rapid growth of Asian economies, and the increasing volumes of international trade, global logistics management in business operations has become more important than ever. In this regard, transportation is becoming a more strategic business function because transport costs account for a larger percentage of the cost of goods sold. There is an increasing interest in reducing transportation costs and increasing route efficiency. Maritime transportation plays a key role in international trade as it represents a low cost transportation mode for high volume and long-distance shipments, being far less expensive than airplane transportation. Hence, maritime transportation is responsible for the majority of long-distance shipments in terms of volume. According to the review of maritime transport by UNCTAD [1], more than seven million tons of goods are carried by ship annually. Some illustrative statistics are provided in [2]. General shipping industry statistics are available in publications by the Institute of Shipping Economics and Logistics (http://www.isl.org/) and the Astrup Fearnley Group (http://www.isl.orgwww.fearnley.com/).

Optimizing maritime transportation systems involves several types of decisions (strategic, tactical, and operational). The strategic decisions include network design (configuration of the routes and their frequencies) and fleet and ship size determination. Tactical decisions include routing and scheduling of ships either for liner, tramp, or industrial shipping. Operational decisions refer to day-to-day decisions which may be aided by the design of on-board advisory systems that increase a vessel's operability and performance.

In this work, we consider a tactical problem consisting of routing and scheduling a heterogeneous tramp fleet. Gatica and Miranda [3] propose a network based model in which time windows for picking and delivering cargoes are discretized and the model is solved directly by using CPLEX. Authors showed that the loss of optimality due to the discretization approach was not significant. Size of the instances solved by the authors considered up to 50 cargo contracts, a fleet size of up to 9 ships, and a level of discretization of 15 time nodes. As reported in the results section, there were several instances that cannot be solved to optimality, which is proportional to the level of discretization. For the continuous case, only 50% of the instances could be solved, and as long as the level of discretization increases, higher number of instances can be solved to optimality.

Difficulty in solving real-life sized problems motivated us to propose, as an extension of this previous work, a heuristic procedure based on a variable neighborhood search to efficiently solve the problem for larger size instances. In order to evaluate the performance of the procedure, we create a set of instances using the same instance generation procedure used in [3] and compare the performance of the heuristic procedures against the results obtained by using CPLEX.

The rest of the paper is organized as follows. Section 1 provides a brief review of related literature and Section 2 presents background of the problem. Section 3 presents the problem description and introduces the notation and the mathematical formulation provided in [3]. Section 4 provides algorithmic and implementation details. Section 5 presents computational experimentation and finally conclusions and recommendations for further research are provided in Section 6.

## 2. Background

The problem addressed in this paper is related to the general traveling salesman problem (TSP), the vehicle routing problem (VRP), and, especially, to a variation of the VRP: the vehicle routing problem with time windows (VRPTW). TSP is a problem based on a salesman who must visit $n$ clients and return to the initial place of departure. The objective is to visit all clients without passing through the ones previously visited [4]. VRP is considered as a generalization of TSP where the clients request either delivery or pick up of an amount of cargo. The VRP differs from the TSP problem in the fact that more than one vehicle is needed to deliver the cargoes with an associated cost [5]. We refer the interested reader to the work of Laporte and Osman [6], for a comprehensive review on VRP and VRPTW, and to the work of Ando and Taniguchi [7], which discusses recent issues arising on city logistics and urban freight transport.

Christiansen et al. [8] discuss several differences between ship routing and other vehicle routing problems and justifies the need for research specifically focused on ship routing and scheduling. The most recent review of ship routing and scheduling is presented in [9], in which research on ship routing and scheduling problems during the new millennium is reviewed. Some of the highlights are that the number of papers doubles every decade and that research on liner shipping, marine inventory routing, and optimal speed is leading the research efforts. In [8], a literature review for these problems is provided and perspectives for further research as well as other optimization and decision-support techniques within the shipping industry are discussed. For earlier literature reviews, the reader is referred to [10, 11].

Three general modes of operation for shipping companies can be distinguished: industrial, tramp, and liner [8, 11]. In liner shipping, the ships follow a published schedule with regular itineraries and predetermined routes, frequencies, and port arrivals/departures; it is very similar to a bus line. The tramp shipping company follows the available cargoes similarly to a taxi. A tramp shipping company usually has a fixed amount of contract cargoes that it is committed to carry and tries to maximize the profit from optional cargoes. In industrial shipping, the cargo owner usually controls the ships and aims to ship cargo at a minimal cost. In the simplest cases, industrial fleet operation is similar to tramp shipping. In more general cases, the industrial fleet operation becomes more complex, especially when trips are not prespecified and a supply network has to be determined based on time-dependent supply chain demand functions. The main difference is that industrial shipping is commonly used for a specific type of cargo related to a certain type of industry. Tramp is usually the operation mode to transport liquid and dry commodities or cargo involving a large number of units (e.g., vehicles) and liner shipping is the selected mode to transport containerized cargo which represents the major segment of liner shipping [3].

The main costs in ocean shipping are (1) capital and depreciation costs which are related to the loss of a ship's market value with respect to the initial investment, (2) running costs which are fixed costs such as maintenance, insurance, crew salaries, and overhead costs, among others, and (3) operating costs which are associated with day-to-day operations such as fuel consumption, port and customs expenses, and tolls paid at canals, among others. Fuel consumption has been a relevant subject in the maritime industry as well as for the world's largest navies due to oil price variability and environmental considerations which drive the effort for fuel-efficient navigation. Fuel consumption can be, to a large extent, controlled by navigation speed since it is approximately a cubic function of speed [11].

We base our discussion on the recent works on ship routing and scheduling of tramp fleets. Even when maritime transportation is a part of a supply chain, Christiansen et al. [8] found that little work has been done to integrate the whole supply chain. Later, Flatberg et al. [12] developed another solution approach for solving the problem proposed in [11]; they use an iterative improvement heuristic combined with an LP solver. Fox and Herden [13] describe a MIP model to schedule ships from ammonia processing plants (which convert ammonia into different fertilizer products) to eight ports in Australia. The objective is to minimize freight, discharge, and inventory holding cots while taking into account the inventory, minimum discharge tonnage, and ship capacity constraints. The MIP model is solved by using commercial optimization software. An inventory routing problem similar to [11] but with multiple products was analyzed by Ronen [14] for liquid bulk oil cargo. Considering multiple products adds complexity to the model since it requires separating the shipments planning stage from the ship scheduling stage. The methods used were MIP and heuristics.

Christiansen et al. [8] commented on the lack of research on tramp shipping as compared to industrial shipping. One main reason could be the large number of small operations in the tramp market. The first work to introduce a typical tramp ship scheduling problem was presented by Appelgren [15, 16]. DW decomposition was employed to solve it. Instead of minimizing costs, the model maximizes the actual marginal contribution (excluding fixed costs). Kim and Lee [17] developed a prototype decision-support system for ship scheduling in the bulk trade where the scheduling problem is formulated as a set packing problem with similar constraints as in Appelgren's model. Several authors have developed decision-support systems for companies operating both in tramp and industrial modes [18–20]. A tramp routing and scheduling problem that maximizes the profit from operating the fleet was solved by [21, 22]. Brønmo et al. [23] develop a multistart local search heuristic. Korsvik et al. [24] propose a unified tabu search heuristic, which allows infeasible solutions with respect to ship capacity and time. In contrast to the procedure followed by Brønmo et al. [23], Malliappi et al. [25] present a variable neighborhood search heuristic; the results show that this procedure outperforms the previous heuristics. Recent research efforts [26, 27] were conducted in solving problems where cargo may be split among several ships.

This paper revisits the model in [3] and proposes heuristic procedures for solving large-scale problems. Gatica and Miranda [3] developed a network based model for the routing and scheduling of a heterogeneous tramp fleet that aims to minimize the total operating cost of serving a set of trip cargo contracts considering time window constraints at both the origin and destination of cargoes. To the best of our knowledge, this is the first paper that proposes a discretization of time windows for picking up and delivering cargoes. This characteristic allows for a broad variety of features and practical constraints to be considered such as navigation speed to control fuel consumption.

## 3. Discretized Time-Window Approach for Solving a Ship Scheduling and Routing Problem

In this section, we present the description of the ship scheduling and routing problem with discretized time windows. Section 3.1 presents the details of the discretized modeling approach and the characteristics of the problem. Section 3.2 presents the mathematical model proposed by Gatica and Miranda [3] which is considered in this paper.

*3.1. Problem Description.* We consider the routing and scheduling problem for tramp shipping which has been addressed by Gatica and Miranda [3]. This is one of the most relevant and challenging problems faced by decision makers at shipping companies along with planning and operation of the liner fleets. An important difference between tramp and liner operations is that liner shipping allows a more static and long-run operation planning than tramp shipping. Liner shipping operates under fixed routes. In contrast to liner case, tramp shipping faces a more dynamic demand and changes in contracts and routes. This emphasizes the importance to address this problem and to provide optimization models for designing the routes and schedules of the ships as well as the need of developing algorithms that allows efficiently solving this problem repeatedly and within reasonable computational times due to the dynamism of the tramp shipping.

As is typical in tramp shipping, contracts and their required schedules are known beforehand by ship contractors. Contracts correspond to a single shipment between two ports, both with time windows for loading and unloading the cargo. Each ship can serve one contract at a time. The fleet of ships is nonhomogeneous in terms of capacity, speed, fuel consumption, and, hence, costs. Not all ships can serve all contracts because the specific characteristics required per contract (cargo-ship or port-ship incompatibilities), which conditions the arc set of the model's underlying network. Incompatible cargoes may make it infeasible for the corresponding trips to be done consecutively by the same ship which also conditions the network.

For each sequence of cargoes to be served by a single ship, a ballast or empty trip must take place from the delivery port of each cargo to the origin of the next cargo in the sequence, unless the delivery port and next origin happen to coincide. Costs for the trips are computed based on fuel consumption and the distance travelled as well as some other operational variables and may vary from ship to ship due to the heterogeneity of the fleet. Furthermore, the speed of the ship is a key factor that affects cost. On the other hand, income is fixed for each contract and hence profit is maximized by minimizing total costs. Then, the problem consists of defining the set of cargoes to be served by each ship as well as the times of arrival and departure and waiting times at each port, with the aim of serving all cargoes at minimal total cost.

*3.2. Mathematical Formulation.* We adopt the discretized modeling approach used in [3] for the time windows for picking up and delivering cargoes, assuming that the arrival time to the origin port must occur only at discrete times. The same applies for cargo delivery times. Thus, each contract consists of a discrete set of possible time instants in which cargo may be picked up and/or delivered.

The nodes, indexed by $i = 1, \ldots, N$, of the network represent discrete and feasible starting times for each cargo. For each node $i$, the cargo associated with that node is represented by $n(i)$. The set of nodes for cargo $n(i)$ is represented by $D_i$. Ships are indexed by $k = 1, \ldots, B$. Each arc$(i, j, k)$ represents the service of cargoes $n(i)$ and $n(j)$ consecutively by ship $k$ and is included in the network if both the trips and the ship are compatible. Arc$(i, j, k)$ is included in the network if it is feasible for ship $k$ to begin service of cargo $n(i)$ at the time instance represented by node $i$, which can deliver the cargo within the corresponding time windows, make the ballast trip from the destination port of cargo $n(i)$ to the origin of cargo $n(j)$, and be available to begin service of cargo $n(j)$ at the time instance associated with node $j$.

For each arc, the cost parameter $c_{ijk}$ represents the total minimal cost incurred when the ship delivers cargo $n(i)$ immediately followed by cargo $n(j)$. Costs include travel costs for trip between the ports associated with $n(i)$, waiting times, and the ballast trip to the origin port of $n(j)$. To complete the network, a fictitious node 0 is created to represent the source of all ships. For each ship $k$ and node $i$, if cargo $n(i)$ is compatible with ship $k$, there are an arc$(0, j, k)$ and an arc$(i, 0, k)$. Cost $c_{0ik}$ is calculated based on the real initial position of ship $k$, and cost $c_{0ik}$ represents the minimum total

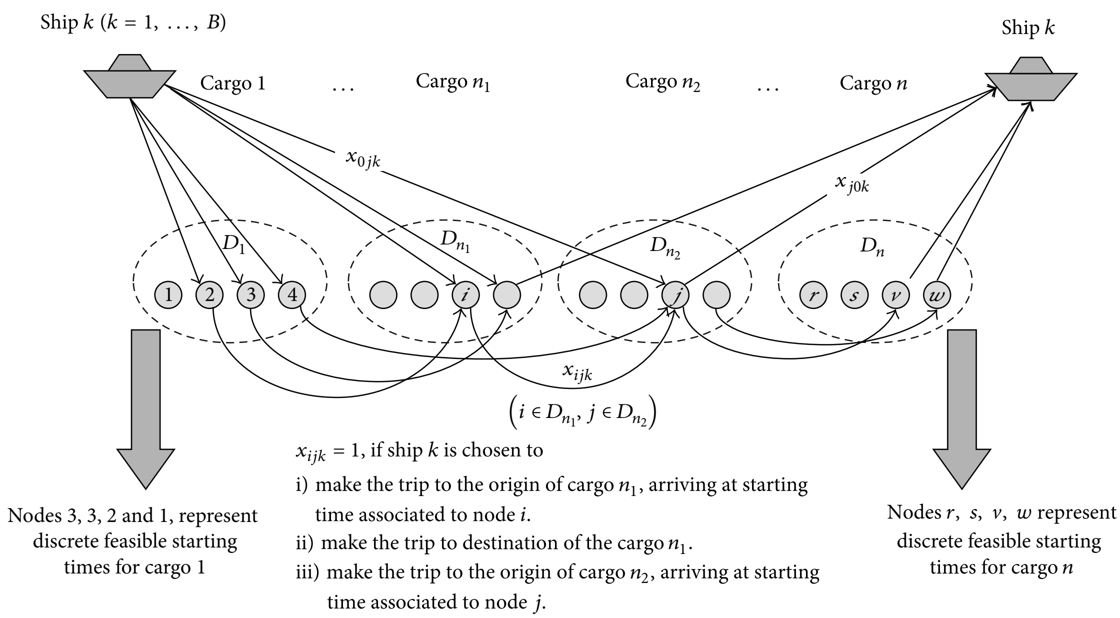



Figure 1: Graph representation of the problem. Source: Gatica and Miranda [3].

cost incurred if ship $k$ serves cargo $n(i)$. It is assumed that ship $k$ will stay at the port of destination of its last assigned cargo.

We can observe a single-ship view of the graph in Figure 1. The segmented ovals group all nodes (feasible starting times) related with the same cargo. Some arcs for a single ship $k$ are drawn, based on the feasible trips that can be selected. Each arc, as previously mentioned, takes into account the ballast trip between destination port of cargo $n(i)$ and the origin of cargo $n(j)$.

The mathematical formulation of the problem from [3] is as follows:

$$\text{Min} \quad \sum_{(i,j,k)\in A} c_{ijk} \cdot x_{ijk} \tag{1}$$

$$\text{s.t.:} \quad \sum_{i\in V/(0,i,k)\in A} x_{0ik} \le 1 \quad k = 1,\ldots,B \tag{2}$$

$$\sum_{(i,j,k)\in A/j\in D_n} x_{ijk} = 1 \quad n = 1,\ldots,N \tag{3}$$

$$\sum_{i\in V/(i,j,k)\in A} x_{ijk} = \sum_{l\in V/(j,l,k)\in A} x_{ijk} \qquad j \in V,\ k = 1,\ldots,B \tag{4}$$

$$x_{ijk} \in \{0,1\} \quad (i,j,k) \in A, \tag{5}$$

where $N$ is number of cargoes or contracts to be served, $V$ is set of nodes in the network, $D_n$ is set of nodes associated with cargo $n$ (i.e., set of possible starting times for trip $n$), $B$ is number of available ships, $A$ is set of arcs in the network, and $c_{ijk}$ is cost of arc$(i, j, k)$. Consider

$$x_{ijk} = \begin{cases} 1 & \text{if arc } (i,j,k) \text{ is selected as part of the solution.} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Selecting arc$(i, j, k)$ as part of the solution ($x_{ijk} = 1$) implies that ship $k$ will serve cargo $n(i)$ and will serve cargo $n(j)$ immediately afterwards. Selecting arc$(0, i, k)$ implies that $n(i)$ is the first cargo to be served by ship $k$, and selecting arc$(i, 0, k)$ implies that $n(i)$ is the last cargo to be served by ship $k$.

The objective function (1) represents the total solution cost. Constraints (2) ensure that each ship is employed at most in one route. A route is defined as a sequence of cargoes to be served. Constraints (3) ensure that, for each cargo $n$, exactly one arc entering set $D_n$ is selected, establishing that each cargo must be served exactly once, by exactly one ship, which begins service at exactly one of the nodes or time instants in the discretized time window for cargo pick up. For nodes other than the central fictitious node, constraints (4) state that if an entering arc is selected, a leaving arc must also be selected and that both arcs must be associated with the same ship. For the fictitious node, this constraint states that if a leaving arc associated with ship $k$ is selected, then an entering arc associated with the same ship must also be selected (i.e., if a ship exits the node), and then it must return to it. Arcs leaving the fictitious node represent the ships that are, in fact, used in the solution.

## 4. Proposed Methodology

There are several contributions related to ship routing and scheduling problems in the literature and several mathematical models have been proposed to optimize related decisions. In addition, diverse solution approaches based on either metaheuristics or mathematical programming methods have been developed. Heuristic procedures are frequently used when exhaustive enumeration and/or optimal solution methods are impractical.

The difficulty of the ship routing and scheduling problem with discretized time windows motivated us to propose a heuristic procedure based on a Variable Neighborhood Search (VNS) metaheuristic structure. In VNS, the basic idea is to explore different vicinities in a systematic way. The majority of local search algorithms use a single neighborhood. VNS is based on three basic concepts: (1) a local minimum with respect to a structure of vicinity is not related to another local minimum with respect to another structure of vicinity; (2) a global minimum is a local minimum with respect to all possible structures of vicinity; and (3) local minima with respect to one or more structures of vicinity are close to each other [21].

Originality of the proposed algorithm relies on the adaptation of a local search heuristic to a routing and scheduling problem that is very challenging and recent, that is, the routing and scheduling of a tramp fleet with variable speed and discretized time windows (Gatica and Miranda [3]) which has not been extensively addressed in the literature (see Christiansen et al. [8]). The proposed heuristic procedure provides a solution to the routing and scheduling problem of a tramp fleet with discretized time windows and variable speed in a more reasonable computational effort which enhances the usability of the proposed method to address large scale instances faster than using CPLEX (as numerical results in Chapter 5 display).

The proposed algorithm consists of two main stages. The first stage searches for a feasible solution that defines a route for each ship. The second stage seeks to improve the initial solution by a local search procedure similar to a VNS mechanism. For this, different neighborhood structures were defined which are sequentially applied in a steepest descent fashion. A solution to the problem consists of a route for each ship in which the route is defined as the set of contracts to be performed by the ship as well as the departure and arrival time at each port (as allowed by the discretized times). Pseudocode 1 presents the pseudocode of the general procedure.

*4.1. Stage 1. Construction of an Initial Feasible Solution.* As shown in Pseudocode 1, the first stage refers to the construction of an initial feasible solution. For this, we employ a greedy procedure. In order to describe this procedure, we introduce the nomenclature described in Table 1, in which each node represents an allowable and discrete time instant in which cargo can be picked up or delivered. Each contract consists of a group of nodes in which the cargo is able to be picked up at an origin and delivered at a destination. Arcs joining nodes represent the associated cost of the trip, based on the distance between the origin-destination pair and the speed required to arrive at the correct time instant at the destination.

The *Construction()* procedure analyzes each contract with the aim of assigning it to a ship. For this, a sorted list of contracts based on their due dates is formed with the earliest due date contract at the top of the list. The ships are also sorted on a list. Initially, the ships are in a random order. The first iteration of the procedure begins by selecting the first contract of the list and the first ship on the list in order to analyze if it is possible to assign the contract to the ship at the earliest time instant in the discrete set of time instances in the time window of the contract. If this is possible, the contract is assigned to the ship and the ship is placed at the last position of the list of ships. Otherwise, we select the next ship of the list and repeat the same procedure until the contract is assigned to a ship in the earliest possible time instant. In each iteration, the ship to which the contract is assigned is placed in the last position of the list of ships. Therefore, the greedy function of this procedure is based on prioritizing the earliest due date contracts and seeking to assign each contract at the earliest possible instant to a ship.

Table 1: Overview of graphical nomenclature.

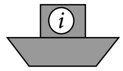
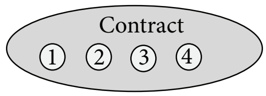
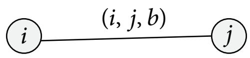
| Graphical nomenclature: | Represents |
|---|---|
|   | Ship $i$. |
|   | Node or instant of time. |
|   | Contract with the corresponding discrete feasible starting times for cargo of the trip. Consider the illustration that the node 1 is the earliest time instant and 4 is the latest time instant. |
|   | Arc representing a ship $b$ that departs at time instant $i$ from the origin port and arrives at time instant $j$ at the destination port. |

Figure 2 illustrates the procedure. In the example, we consider four ships and twelve contracts. The first contract is assigned in its earliest time window to the first ship. The process is repeated for contracts 2 to 5. However, when we analyze contract 6 with the corresponding sequential ship 2, we realize that it is not possible to assign the contract in its earliest time instant to ship 2, so ship 3 must be considered. Given that it is possible to assign the contract to ship 3, this is assigned. Then, we consider contract 7. At the top of the list of ships we still have ship 2, so we explore the possibility to assign it to contract 7 in its earliest time instant. Given that this is possible, we assign the contract. The procedure is repeated until all contracts are assigned to ship at some time instants, always striving to assign the contracts as early as possible. As can be seen in Figure 2, there are some contracts that get assigned to later time instants because it was not possible to assign them to any ship at earlier time instants.

*4.2. Stage 2. Local Search Procedure.* Once a feasible initial solution has been constructed, a local search procedure is applied in order to improve the solution. For this, we propose to explore iteratively and in a steepest descent fashion, using four neighborhood structures. If no improvement of the solution is attained, then we consider two additional alternative neighborhood structures that will be also explored in a steepest descent fashion. Pseudocode 2 presents the pseudocode of the general procedure.

*4.2.1. ImproveRoute().* This neighborhood of solution $x$ consists of the set of solutions that results from exploring all feasible combinations of arcs that connect two contracts in

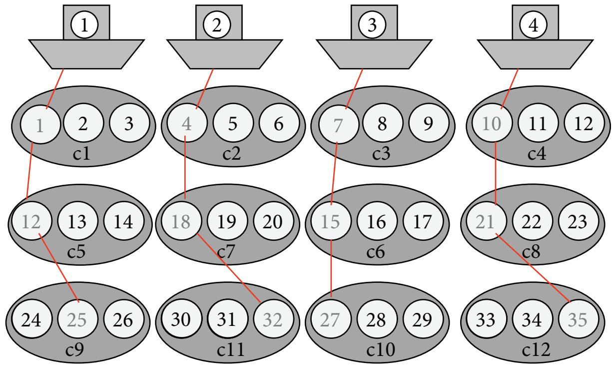



Figure 2: Example of *Construction()* procedure.

Ship route Procedure
Input: $P$ := a problem instance $(N, V, D_n, B, A, c_{ijk})$
Output: $x^*$ := Route for each ship $k$,
or empty set with no feasible solution.
$x \leftarrow$ Construction();
Set $f^* = f(x)$; $x^* = x$
Do until a stopping condition is met
$x \leftarrow$ Local-Search();
If $f(x) < f^*$ then
$f^* = f(x)$ and $x^* = x$;
End Do
Return $x^*$

Pseudocode 1: Pseudocode of the solution procedure.

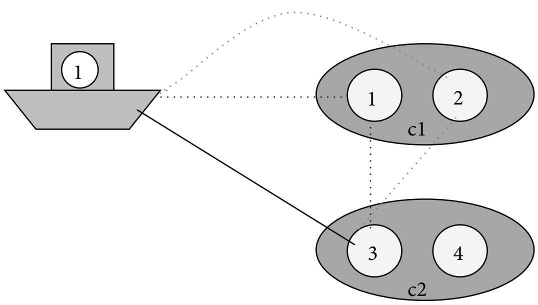



Figure 3: Illustration of an iteration of the *ImproveRoute()* procedure.

the route of a ship, selecting the pair of arcs with lowest cost. The procedure aims to improve solutions based on an analysis of the time windows of each contract, respecting the contracts assigned to the route of a ship. Figure 3 shows different options for a ship (each dotted line color corresponds to an option).

*4.2.2. InsertionContractsN().* This neighborhood structure consists of the set of solutions that results from moving a contract from a route to insert it into another route. For this, the procedure considers a pair of ship routes and evaluates both the active nodes and nonactive nodes (not included in the initial solution). If moving a contract from one route to another improves the solution reducing total costs, then the move is performed. The order of the contracts is maintained at all times.

Figure 4 illustrates the procedure. The initial solution (black lines) consists of the routes of ship 1 (contract 1 at time instant 2, followed by contract 2 at time instant 3) and ship 2 (contract 3 at time instant 6). The procedure then analyzes the insertion of a contract from the route of ship 1 into the route of ship 2. The yellow arcs are an example of an option that includes nonactive nodes of the initial solution.

As observed in Figure 4, analyzing the insertion of contract 2 into the route of ship 2, it turns out that costs are reduced if contract 2 is inserted into the route of ship 2 as shown by the red lines. The resulting route for ship 2 selects contract 2 at time instant 3 followed by contract 3 at time instant 6.

*4.2.3. InterchangeContractsN().* This neighborhood structure consists of the set of solutions that results from interchanging contracts between two routes of ships considering both the active nodes and nonactive nodes (those nodes that are part of the current solution as well as those that are not) and respecting the initial order of contracts in the routes. The exchange is performed only if the new configuration provides lower costs.

Figure 5 illustrates this method. Consider the initial solution presented in part (a) of the figure and assume that we will evaluate the exchange of contracts 1 and 3 (c1 and c3). As can be seen in part (b) of the figure, we add arcs to get a new solution (red, blue, pink, and yellow dotted lines). Black dotted lines correspond to the initial solution and the black line indicates that contract 2 should be performed by ship 1. If the new configuration provides lower costs, the contracts are interchanged and a new solution is obtained. Suppose this is the case for solution found with pink and yellow dotted arcs, as shown in (c).

*4.2.4. TwoOptUpward().* This neighborhood of solution $x$ consists of the set of solutions that results from crossing over a pair of routes, based on a variation of the 2 Opt local search algorithm proposed by [22]. The procedure considers a pair of arcs to be crossed and the upward part of the each

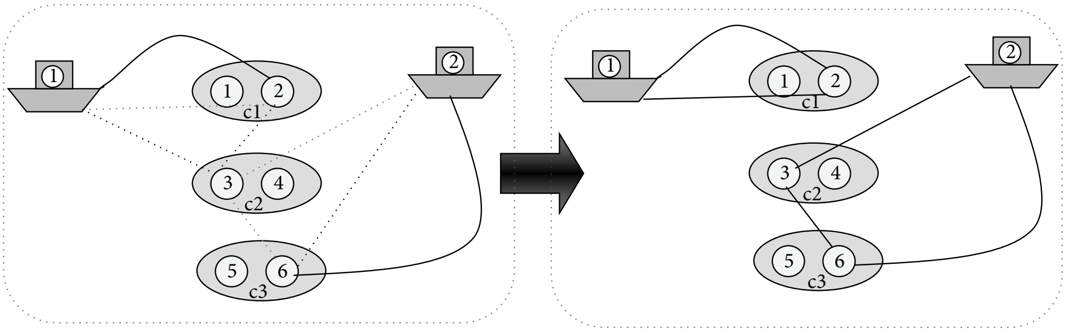
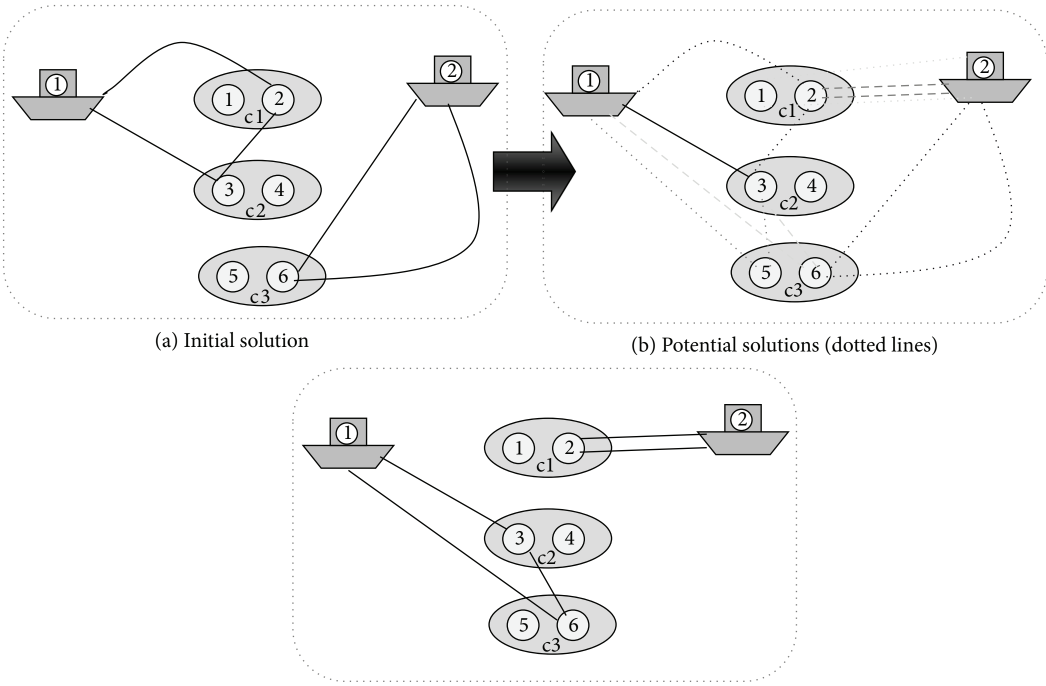


Figure 4: Illustration of *InsertionContractsN()* procedure.



(a) Initial solution

(b) Potential solutions (dotted lines)

(c) Best solution

Figure 5: Illustration of *InterchangeContractsN()* procedure.

route is swapped. The procedure differs with respect to the 2 Opt procedure in that segments of several contracts are exchanged and not only the individual route is improved. Figure 6 illustrates the procedure.

*4.2.5. InsertionContractsN2().* This neighborhood of solution $x$ is a variant of the *InsertionContractsN()* but differs in that when a contract is inserted between a pair of contracts of another route, for the previous contract to the one inserted, we select the earliest possible node for its departure and for the following contract in the route to the one inserted, we select the latest possible node for arriving. The insertion is performed only if lower costs are obtained. The procedure is illustrated in Figure 7.

*4.2.6. InterchangeContracts_S().* This neighborhood structure is a simplification of the *InterchangeContractsN()* procedure in which solutions that result from all possible pair of contracts to be exchanged between two ship routes are evaluated, considering only active nodes (those that are currently part of the solution). The exchange is performed only if the new configuration provides lower costs.

*4.2.7. InterchangeContracts_C().* This neighborhood structure is a variant of previous one (Interchange Contracts-S) in which a pair of contracts is exchanged but instead of searching among a pair of ships, the search is performed on a contracts sequence.

## 5. Computational Experimentation

This section describes the test instances generation and the computational results. The heuristic procedure was implemented in JAVA SE 6 and numerical experimentation was performed to test its performance. We generate a set of instances of different sizes. Previous instances solved in [3]

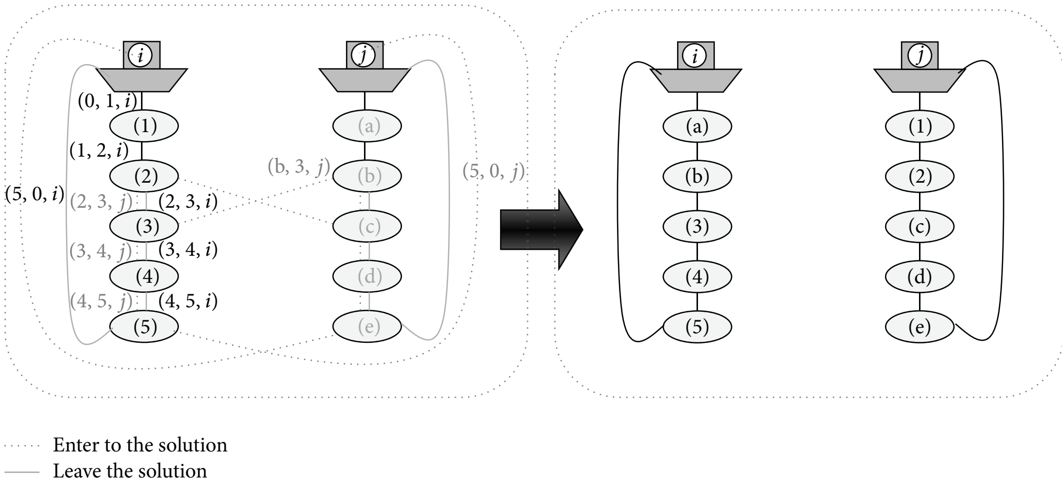



Figure 6: Illustration of an iteration of the *TwoOptUpward()* procedure.

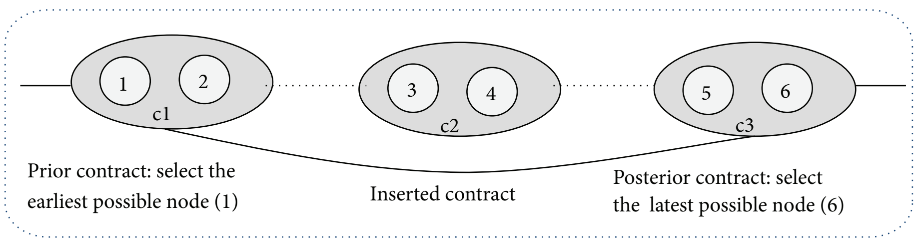



Figure 7: Pseudocode of the local search procedure.

were not available, so that new instances were generated using the test instances generator code developed by [3]. For comparison purposes, all instances were also solved by CPLEX 11.1. Numerical experiments were performed using a 2.00 GHz Pentium processor with 2 GB of RAM running under Windows XP.

*5.1. Instance Generation.* An instance of the problem consists of a list of contracts and the corresponding set of nodes for arrival and departure arrivals, as well as a set of arcs for all the feasible trips that could be performed. Each arc possesses an initial node, a final node, a ship number, and an associated cost. As it was previously mentioned, we employed the test instances generator developed by [3] which is composed of a real based database containing potential contracts in a port-port matrix, including distances of trip estimated to each potential contract.

The database contains 400 potential cargo contracts, a port-port matrix which contains distances of trips estimated for ach potential cargoes among 87 different ports. For each potential cargo, a pair of ports was chosen with real place of arrival which is frequently used in practice. Information regarding the necessity of going through a channel for each trip is used as well as the channel fees. In order to generate an instance, a subset of the cargoes from the database is randomly selected, according to the instance size. For the generation of random numbers, 45 different seeds are employed. The application is coded in JAVA SE 6. It is important to recall that the instances generated should guarantee that it is possible to obtain a feasible solution.

Eighteen groups of instances were generated, each group composed of a combination of ships, time window nodes, and contracts. The set of discrete time window consists of 3, 6, or 15 nodes. The number of ships was varied over the values of 4, 5, 7, and 9. The number of contracts was varied over the values of 30, 40, and 50. Each group contains 15 different instances. Instance sizes consider between 14,000 and 2,000,000 arcs. In total, we generated $3 \times 4 \times 3 \times 15 = 540$ instances.

*5.2. Results and Discussion.* In this section, a comparison of the results obtained by using CPLEX and the heuristic is presented. Tables 1 and 2 present the results according to the instance groups (18 groups generated with 15 instances each). Stopping rules consider a limit time of 7200 seconds for CPLEX. For the heuristic no limit time was set, considering only a maximum number of iterations without any improvement in the solution as stopping criterion based on an epsilon which was defined in terms of the instance size. For comparison purposes, instance sizes in which CPLEX could find at least a feasible solution are considered which are of similar size as those solved in [3]. Hence, for each instance solved, we compare the results obtained by CPLEX and the heuristic and estimate a gap with respect to the best integer solution reported by CPLEX (which for some

Table 2: Summary of results of instances of 30 contracts.

| Instance | | | CPLEX | | | | Heuristic | | |
|---|---|---|---|---|---|---|---|---|---|
| Contracts | Ships | Nodes per window | Optimum found | Only feasible solution | No solution | Average time (sec) | Solution not found | Average time (sec) | GAP O.F. (%) average |
| 30 | 4 | 3 | 13 | 0 | 2 | 1 | 3 | 4.80 | 4.79 |
| 30 | 4 | 6 | 13 | 0 | 2 | 4.84 | 2 | 15.92 | 5.36 |
| 30 | 4 | 15 | 13 | 0 | 2 | 171.84 | 2 | 93.15 | 5.53 |
| 30 | 5 | 3 | 15 | 0 | 0 | 1.13 | 0 | 7.06 | 4.34 |
| 30 | 5 | 6 | 15 | 0 | 0 | 12.73 | 0 | 26.53 | 5.49 |
| 30 | 5 | 15 | 15 | 0 | 0 | 140.60 | 0 | 151.40 | 5.49 |

```
Local-Search()
Input: x := Initial Feasible solution (Route for each ship k)
Output: x* := Best Solution found (Route for each ship k)
  Do until no further improvement is achieved
    InsertionContractsN()
    ImproveRoute()
  End do
  Do until no further improvement is achieved
    InterchangeContractsN()
    ImproveRoute()
  End do
  Do until no further improvement is achieved
    2OptUpward()
    ImproveRoute()
  End do
  Do until no further improvement is achieved
    InsertionContractsN2()
    ImproveRoute()
  End do
  If no improvement of the initial solution was found, then
    InterchangeContracts_S()
    If no improvement was previously found, then
      InterchangeContracts_C()
    End if
  End If
  Return x*
```

Pseudocode 2: Pseudocode of the Local Search Procedure.

instances corresponds to the optimal solution) as described by (7) in which $Z$ corresponds to the value obtained by the heuristic. Positive gaps are obtained when CPLEX finds better solutions. Consider

$$\text{variation} = \frac{Z - \text{CPLEX}}{\text{CPLEX}} * 100. \tag{7}$$

Tables 2, 3, and 4 show a summary of all results obtained for those instances of 30, 40, and 50 contracts, respectively. Each instance type is indicated according to its combination of contracts, ships, and nodes which account for 15 replicates of each instance type. The averages of 15 replicates are shown for the execution times of CPLEX and the execution times of the heuristic. Furthermore, average gaps computed according to (7) are also presented, considering only those cases in which at least a feasible solution was obtained by CPLEX or by the heuristic. The tables indicate, for each instance type, the number of instances in which an optimum solution was found, the instances in which a feasible (nonoptimal) solution was found, and also those cases in which no feasible solution was obtained by CPLEX. Similarly, for the heuristic procedure, the tables present the number of instances in which an initial solution could not be found and, consequently, could not apply the methods of local search.

Table 2 shows instances of 30 contracts corresponding to the smaller size instances in which the difference on computational times between CPLEX and the heuristic resulted no significant. In terms of the quality of solutions, average gaps of the solution found by the heuristic with respect to CPLEX is less than 5.5%.

Instances of 40 contracts are medium size instances and we can observe from Table 3 a more significant difference between computational times of CPLEX and the heuristic. In terms of the gaps found by the heuristic solutions with respect to CPLEX, the maximum average gap corresponds to 8.11%. In most of the instances, CPLEX found an optimal solution or at least a feasible solution.

Table 4 shows instances of 50 contracts correspond to the biggest size instances, and computational times of CPLEX and the heuristic present more significant differences. In terms of average gaps, we observe even lower gaps with respect to the instances of 40 contracts, where the maximum average gap corresponds to 6.69%. It is noteworthy that the heuristic improves the execution times for large problems, which indicates that it can be employed for solving larger size instances. On the other hand, we can observe a good performance of the procedure with relatively small gaps with respect to the optimum solutions found by CPLEX, with a maximum value of 8.11%, which corresponds to a medium size instance (40 contracts) and does not increment proportionally to instance size.

Tables 5, 6, and 7 present results classified according to the number of ships and nodes per window for the instances of 30, 40, and 50 contracts, respectively. The tables present the percentage of instances in which CPLEX found an optimum solution, a feasible solution, and no solution, based on the number of instances ran for each type. Results found by the heuristic procedure are also presented in terms of the percentage of instances in which no solution was found. Average computational times are presented for both CPLEX

Table 3: Summary of results of instances of 40 contracts.

| Instance | | | CPLEX | | | | Heuristic | | |
|---|---|---|---|---|---|---|---|---|---|
| Contracts | Ships | Nodes per window | Optimum found | Only feasible solution | No solution | Average time (sec) | Solution not found | Average time (sec) | GAP O.F. (%) average |
| 40 | 5 | 3 | 14 | 0 | 1 | 3.85 | 6 | 21.55 | 5.00 |
| 40 | 5 | 6 | 15 | 0 | 0 | 42 | 3 | 82.50 | 8.11 |
| 40 | 5 | 15 | 15 | 0 | 0 | 719.40 | 3 | 442.58 | 7.49 |
| 40 | 7 | 3 | 15 | 0 | 0 | 8.66 | 0 | 33.93 | 6.27 |
| 40 | 7 | 6 | 15 | 0 | 0 | 447.53 | 0 | 137.13 | 7.38 |
| 40 | 7 | 15 | 12 | 3 | 0 | 1870.86 | 0 | 645.53 | 7.76 |

Table 4: Summary of results of instances of 50 Contracts.

| Instance | | | CPLEX | | | | Heuristic | | |
|---|---|---|---|---|---|---|---|---|---|
| Contracts | Ships | Nodes per window | Optimum Found | Only Feasible Solution | No Solution | Average Time (sec) | Solution not Found | Average Time (sec) | GAP O.F. (%) Average |
| 50 | 7 | 3 | 14 | 0 | 1 | 18.71 | 2 | 109.30 | 5.95 |
| 50 | 7 | 6 | 13 | 0 | 2 | 183.53 | 4 | 358.45 | 5.35 |
| 50 | 7 | 15 | 8 | 7 | 1 | 4291.85 | 2 | 1551.38 | 5.87 |
| 50 | 9 | 3 | 15 | 0 | 0 | 23.80 | 0 | 134.06 | 6.69 |
| 50 | 9 | 6 | 15 | 0 | 0 | 371.53 | 0 | 558.40 | 6.53 |
| 50 | 9 | 15 | 7 | 8 | 0 | 4450.20 | 0 | 2795.53 | 6.09 |

Table 5: Results per contracts.

| Type of instance | CPLEX | | | | Heuristic | | | |
|---|---|---|---|---|---|---|---|---|
| Contracts | Optimum found | Only feasible solution | No solution | Average time (sec) | Solution not found | Average time (sec) | GAP % time average | GAP % OF average |
| 30 | 93.33% | 0.00% | 6.67% | 55.36 | 7.78% | 49.81 | −0.1001 | 5.17 |
| 40 | 95.56% | 3.33% | 1.11% | 515.38 | 13.33% | 227.20 | −0.5591 | 7.00 |
| 50 | 80.00% | 16.67% | 4.44% | 1556.60 | 8.89% | 917.85 | −0.4103 | 6.08 |

and the heuristic. Average gaps of the heuristic with respect to CPLEX are also presented for each type of instance.

As observed in Table 5, computational times for CPLEX and the heuristic increase as long as the number of contracts also increase. However, not necessarily the number of instances in which CPLEX cannot find an optimal solution increases proportionally as the number of contracts increases, given that for 40 contracts the percentage of optimal solutions found by CPLEX is higher than for 30 contracts, but for 50 contracts it is observed significant reduction on the number of optimal solutions found by CPLEX. In the case of the heuristic, similar results are found in which the percentage of cases in which no feasible solution is found decreases for the 50 contracts instances. For the 50 contracts instances, computational times of the heuristic procedure are significantly lower than CPLEX times.

As observed in Table 6, the number of ships does not significantly increase the difficulty of the instance. Similar results in terms of the number of optimal solutions are found by CPLEX for most of the cases. For the heuristic procedure, very similar gaps are obtained for all the instances which, on average, are about 6%. Average times increase for both CPLEX and the heuristic for the instances with more ships which was expected.

As shown in Table 7, as the number of nodes increases, it becomes more difficult to obtain exact solutions by CPLEX and computational times increase significantly and, in this case, the heuristic performs better. On the other hand, gaps of the heuristic with respect to CPLEX do not increase proportionally with respect to the difficulty of the problem and are about 6%.

Provided that in some instances CPLEX found more efficient solutions than the heuristic procedure; Tables 8 and 9 present an analysis of the results to determine which the cases are in which the proposed heuristic is efficient. Table 8 presents a comparison of computational times between the heuristic and CPLEX considering the number of nodes per window and the number of contracts. An index is computed as indicated at (8). When the index is less than 1, the heuristic achieves better results. Consider

$$\text{Index} = \frac{\text{TimeHeuristic}}{\text{TimeCPLEX}}. \qquad (8)$$

Table 6: Results per ships.

| Type of instance | CPLEX | | | | Heuristic | | | |
|---|---|---|---|---|---|---|---|---|
| Ships | Optimum found | Only feasible solution | No solution | Average time (sec) | Solution not found | Average time (sec) | GAP % time average | GAP % OF average |
| 4 | 86.67% | 0.00% | 13.33% | 59.23 | 15.56% | 37.96 | −0.3591 | 5.23 |
| 5 | 98.89% | 0.00% | 1.11% | 153.29 | 13.33% | 121.94 | −0.2045 | 5.99 |
| 7 | 85.56% | 11.11% | 4.44% | 1136.86 | 8.89% | 472.62 | −0.5842 | 6.43 |
| 9 | 82.22% | 17.78% | 0.00% | 1615.18 | 0.00% | 1162.66 | −0.2801 | 6.44 |

Table 7: Results per nodes per window.

| Type of instance | CPLEX | | | | Heuristic | | | |
|---|---|---|---|---|---|---|---|---|
| Nodes per window | Optimum found | Only feasible solution | No solution | Average time (sec) | Solution not found | Average time (sec) | GAP % time average | GAP % of average |
| 3 | 97.33% | 0.00% | 2.67% | 9.525 | 10.67% | 51.78 | 4.4365 | 5.50 |
| 6 | 95.56% | 0.00% | 4.44% | 177.03 | 10.00% | 196.49 | 0.1099 | 6.370 |
| 15 | 77.78% | 20.00% | 3.33% | 1940.79 | 7.78% | 946.60 | −0.5122 | 6.371 |

Table 8: Analysis of computational times.

| Contracts/nodes | 3 | 6 | 15 |
|---|---|---|---|
| 30 | 5.568075117 | 2.416050085 | 0.78271028 |
| 40 | 4.434852118 | 0.448654832 | 0.420077521 |
| 50 | 5.724770642 | 1.651803409 | 0.497241494 |

Table 9: Analysis of quality of the heuristic.

| Contracts/nodes | 3 | 6 | 15 |
|---|---|---|---|
| 30 | 4.565 | 5.425 | 5.51 |
| 40 | 5.635 | 7.745 | 7.625 |
| 50 | 6.32 | 5.94 | 5.98 |

As observed in Table 8, the heuristic does not present an advantage for the smallest size instances (those with 3 nodes and for the 3 variations of contracts). For the medium size instances (those with 6 number of nodes), results are not conclusive (for some cases the heuristic performs better and for other it does not). However, for the bigger size instances (those with 15 nodes), the heuristic presents better results independent of the number of contracts. Hence, the heuristic is more effective (in terms of computational effort) when the instance contains more nodes. As observed in Table 8, the number of contracts does not increase significantly the complexity of the instance. This can be attributed to the fact that when the number of contracts is increased, the number of ships also increases. Based on the results, when the number of nodes (and consequently the complexity of the problem) is increased, the heuristic is a more attractive option. It is worth noting that the discretization of time windows was the approach followed to tackle a continuous time problem in this paper. As more precision is demanded (i.e., more time windows or nodes are needed), the heuristic results in a more attractive option than using CPLEX.

Table 9 presents an analysis of the quality of the heuristic in terms of the GAP obtained with respect to CPLEX considering the number of contracts and nodes per window. The average gaps of each type of instance are shown. As observed in the table, gaps do not increase with respect to the size of the instance and are relatively similar for all the instance types, with an average of about 6% which is a good feature of the proposed heuristic.

## 6. Conclusions and Further Research

We present a heuristic based on variable neighborhood search heuristic to solve a routing and scheduling problem for tramp shipping operations that are modeled adopting a time based discretization as proposed in [3]. This heuristic approach is an alternative method for solving large instances that CPLEX cannot solve efficiently or even find a feasible solution in reasonable times. Several special features such as navigation speed and time windows are introduced as network parameters in the instances, without increasing the complexity of the heuristic.

Numerical results show that the proposed solution approach requires reasonable computational times with reasonable gaps with respect to the solution found by CPLEX which in general were less than 8% and were about 6% in average. Instances tested in this work represent the size of real instances for a medium size shipping company for tramp mode.

As observed in Section 5, numerical results show that the number of nodes per window is the main parameter that affects the difficulty of the instance. It was observed that the heuristic procedure works better for bigger size instances; hence, if more precision is demanded for the problem (more nodes), then the heuristic represents an efficient method. Furthermore, numerical results indicate that the heuristic presents a similar performance for all the instances and gaps do not increase with respect to the instance size, which is an important element of the heuristic and it may be expected a good and robust performance for bigger size instances in which no comparison with respect to CPLEX is possible.

The usability and applicability that the proposed heuristic provides to solve the routing and scheduling problem of a tramp fleet with discretized time windows and variable speed allow addressing instances of bigger size with less computational times than when using CPLEX. Due to comparison purposes, only instances in which CPLEX is able to find a solution were solved but it is expected that bigger size instances may be able to be solved in reasonable times without decreasing the quality of the solutions as it was observed in the numerical results where the quality of the heuristic did not decrease when the size of the instance increased.

Other features that can be easily incorporated into the heuristic are, for instance, congestion at the ports or priorities of the cargoes, which for the case of the mathematical model may increase its complexity. In particular, for the case of variable navigation speed, a very complex problem is generated even for median size instances, resulting in high computational times even for instances or reduced size. This can be observed in the discretization approach. An instance of 50 cargoes and 15 window times generates $50 \times 15 = 750$ nodes.

The approach model and heuristic presented in this paper provide a decision-support tool for helping shipping companies in the design of the routes and schedules of their tramp fleets. This problem is frequently faced by the industry of international trade which presents an increasing trend in the current global environment and maritime shipping represents the bigger participation among the different transportation modes. Maritime industry presents a high dynamism and variability on its operations; thus, decision-support tools to improve their operations are extremely important. This model and solution procedure provide a mechanism to efficiently plan the routes of the fleet in order to minimize costs associated with the consumption of fuel (which can be, to a large extent, controlled by navigation speed) and to reduce lead times. In this regard, both governments and private industry may potentially benefit from the reduction of logistics costs in which the maritime fleet cost represents an important percentage of the total logistical costs.

For further research we recommend generating instances in which the navigation speeds and the discretized time windows may be determined through the local search instead of including fixed values within the network model. Additionally, mathematical programming techniques for solving the algorithm could be explored. For instance, Branch and Cut algorithms have been widely used for solving vehicle routing problems and may prove to be useful for this problem as well. We also propose applying the model and solution procedure in some real-world case studies in order to measure real improvements with respect to the current operations. Another avenue of research could involve attempting to apply similar models with discretized time windows to liner shipping operations, where the main difference is that usually no ballast or empty trips are required as in tramp shipping.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This research was supported by the Office of the Vice President for Research of the University of Texas at San Antonio through the Tenure-Track Research Grant. The authors would like to thank Oliver Riquelme and Axel Wilson, who worked on the problem addressed in this paper to obtain their bachelor degree in Industrial Engineering at the Catholic University of Valparaiso, making a valuable contribution to this research.

## References

[1] UNCTAD, "Review of Maritime Transport," U.N. Publication, 2009.

[2] R. Agarwal and Ö. Ergun, "Ship scheduling and network design for cargo routing in liner shipping," *Transportation Science*, vol. 42, no. 2, pp. 175–196, 2008.

[3] R. A. Gatica and P. A. Miranda, "Special issue on Latin-American research: a time based discretization approach for ship routing and scheduling with variable speed," *Networks and Spatial Economics*, vol. 11, no. 3, pp. 465–485, 2011.

[4] H. P. Williams, "Models for solving the travelling salesman problem," in *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, R. Barták and M. Milano, Eds., vol. 3524, pp. 17–18, Springer, Berlin, Germany, 2005.

[5] G. Laporte, "The vehicle routing problem: an overview of exact and approximate algorithms," *European Journal of Operational Research*, vol. 59, no. 3, pp. 345–358, 1992.

[6] G. Laporte and I. H. Osman, "Routing problems: a bibliography," *Annals of Operations Research*, vol. 61, no. 1, pp. 227–262, 1995.

[7] N. Ando and E. Taniguchi, "Travel time reliability in vehicle routing and scheduling with time windows," *Networks and Spatial Economics*, vol. 6, no. 3-4, pp. 293–311, 2006.

[8] M. Christiansen, K. Fagerholt, and D. Ronen, "Ship routing and scheduling: status and perspectives," *Transportation Science*, vol. 38, no. 1, pp. 1–18, 2004.

[9] M. Christiansen, K. Fagerholt, B. Nygreen, and D. Ronen, "Ship routing and scheduling in the new millennium," *European Journal of Operational Research*, vol. 228, no. 3, pp. 467–483, 2012.

[10] D. Ronen, "Cargo ships routing and scheduling: survey of models and problems," *European Journal of Operational Research*, vol. 12, no. 2, pp. 119–126, 1983.

[11] D. Ronen, "Ship scheduling: the last decade," *European Journal of Operational Research*, vol. 71, no. 3, pp. 325–333, 1993.

[12] T. Flatberg, J. Havardtun, O. Kloster, and A. Lokketangen, "Combining exact and heuristic methods for solving a Vessel Routing Problem with inventory constraint and time windows," *Ricerca Operativa*, vol. 29, no. 91, pp. 55–68, 2000.

[13] M. Fox and D. Herden, "Ship scheduling of fertilizer products," *OR Insight*, vol. 12, no. 2, pp. 21–28, 1999.

[14] D. Ronen, "Marine inventory routing: shipments planning," *Journal of the Operational Research Society*, vol. 53, no. 1, pp. 108–114, 2002.

[15] L. H. Appelgren, "A column generation algorithm for a ship scheduling problem," *Transportation Science*, vol. 3, no. 1, pp. 53–68, 1969.

[16] L. H. Appelgren, "Integer programming methods for a vessel scheduling problem," *Transportation Science*, vol. 5, no. 1, pp. 64–78, 1971.

[17] S.-H. Kim and K.-K. Lee, "An optimization-based decision support system for ship scheduling," *Computers and Industrial Engineering*, vol. 33, no. 3-4, pp. 689–692, 1997.

[18] D. O. Bausch, G. G. Brown, and D. Ronen, "Scheduling short-term marine tansport of bulk products," *Maritime Policy and Management*, vol. 25, no. 4, pp. 335–348, 1998.

[19] H. D. Sherali, S. M. Al-Yakoob, and M. M. Hassan, "Fleet management models and algorithms for an oil-tanker routing and scheduling problem," *IIE Transactions*, vol. 31, no. 5, pp. 395–406, 1999.

[20] K. Fagerholt, "A computer-based decision support system for vessel fleet scheduling—experience and future research," *Decision Support Systems*, vol. 37, no. 1, pp. 35–47, 2004.

[21] N. Mladenović and P. Hansen, "Variable neighborhood search," *Computers and Operations Research*, vol. 24, no. 11, pp. 1097–1100, 1997.

[22] G. A. Croes, "A method for solving traveling salesman problems," *Operations Research*, vol. 6, pp. 791–812, 1958.

[23] G. Brønmo, M. Christiansen, K. Fagerholt, and B. Nygreen, "A multi-start local search heuristic for ship scheduling-a computational study," *Computers & Operations Research*, vol. 34, no. 3, pp. 900–917, 2007.

[24] J. E. Korsvik, K. Fagerholt, and G. Laporte, "A tabu search heuristic for ship routing and scheduling," *Journal of the Operational Research Society*, vol. 61, no. 4, pp. 594–603, 2010.

[25] F. Malliappi, J. A. Bennell, and C. N. Potts, "A variable neighborhood search heuristic for tramp ship scheduling," in *Computational Logistics*, vol. 6971 of *Lecture Notes in Computer Science*, pp. 273–285, 2011.

[26] H. Andersson, M. Christiansen, and K. Fagerholt, "Maritime pickup and delivery problems with split loads," *INFOR*, vol. 49, no. 2, pp. 79–91, 2011.

[27] J. E. Korsvik, K. Fagerholt, and G. Laporte, "A large neighbourhood search heuristic for ship routing and scheduling with split loads," *Computers & Operations Research*, vol. 38, no. 2, pp. 474–483, 2011.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 218913, 17 pages
http://dx.doi.org/10.1155/2014/218913

*Research Article*

# Design and Optimization of Capacitated Supply Chain Networks Including Quality Measures

**Krystel K. Castillo-Villar,[1] Neale R. Smith,[2] and José F. Herbert-Acero[3]**

[1] *Department of Mechanical Engineering, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA*
[2] *Quality and Manufacturing Center, Instituto Tecnológico y de Estudios Superiores de Monterrey, Eugenio Garza Sada 2501 Sur, 64849 Monterrey, NL, Mexico*
[3] *Department of Physics, Instituto Tecnológico y de Estudios Superiores de Monterrey, Eugenio Garza Sada 2501 Sur, 64849 Monterrey, NL, Mexico*

Correspondence should be addressed to Krystel K. Castillo-Villar; krystel.castillo@utsa.edu

Received 17 November 2013; Accepted 18 February 2014; Published 30 March 2014

Academic Editor: Vyacheslav Kalashnikov



This paper presents (1) a novel capacitated model for supply chain network design which considers manufacturing, distribution, and quality costs (named SCND-COQ model) and (2) five combinatorial optimization methods, based on nonlinear optimization, heuristic, and metaheuristic approaches, which are used to solve realistic instances of practical size. The SCND-COQ model is a mixed-integer nonlinear problem which can be used at a strategic planning level to design a supply chain network that maximizes the total profit subject to meeting an overall quality level of the final product at minimum costs. The SCND-COQ model computes the quality-related costs for the whole supply chain network considering the interdependencies among business entities. The effectiveness of the proposed solution approaches is shown using numerical experiments. These methods allow solving more realistic (capacitated) supply chain network design problems including quality-related costs (inspections, rework, opportunity costs, and others) within a reasonable computational time.

## 1. Introduction

The supply chain (SC) can be understood as the integration of all business entities that work together in order to ensure that the customer receives a product or service at the right time, with the right quality, and at low cost. To achieve this, it is necessary to coordinate all the business entities within a SC. This can be achieved through supply chain management (SCM). This paper addresses one of the problems included in SCM which is supply chain network design (SCND). The SCND aims at selecting the business entities that increase the overall performance of the SC. Cost of Quality (COQ) is a measurement system that translates the implications of poor quality into monetary terms. Although COQ has been applied mostly within enterprises, COQ can be applied as an external measure to integrate these costs into SCND modeling. Several studies have provided models to ensure quality in multistage SC [1]. Das [2] proposed a multistage global SC mathematical model for preventing recall risks. Srivastava [3], who initiates estimating COQ in a SC, measures COQ in monetary terms at selected third-party contract manufacturing sites of a pharmaceutical company. Ramudhin et al. [4] also focus on integrating COQ in the SC. Their seminal study presents a mathematical formulation that integrates known COQ functions into the modeling of a SC network for a single-product, three-echelon system and seeks to minimize the overall operational and quality costs. Ramudhin et al. [4] found that by adding a known and given quadratic COQ function that affects only the suppliers into the objective function results in a difference of approximately 16% in costs and changes the network selection. When COQ is not included, choices made solely on production costs could sacrifice quality and lead to additional quality nonconformance costs or corrective action costs in the subsequent stages of the SC. More recently, Alzaman et al. [5] established a mathematical model, considering an $n$ level bill of materials, that incorporates a known COQ quadratic function based on a defect ratio at all SC nodes. As assumed in Ramudhin et al.'s work, the COQ function is known and is based on Juran's original model [6].

In previous studies, the COQ function, based on percentage of defective units, is assumed to be given. This paper deals with the development of an SCND-COQ model that computes the COQ for a whole SC based on internal decisions within the manufacturing plant, such as fraction defective at the manufacturing plant and error rate at inspection. No previous work has addressed how the COQ functions were obtained while taking internal operational decisions within the SC. Moreover, the proposed SCND-COQ model computes quality costs for the whole SC considering the interdependencies among business entities, whereas previous works have assumed independent COQ functions for each node of the SC.

This study aims to develop a strategic-level model for computing the COQ for a multistage, capacitated supply chain network design (SCND-COQ) problem. The proposed model is an extension of the serial supply chain model (SC-COQ model) developed by Castillo-Villar et al. [7] for which two solution procedures were developed [8]. The problem addressed here is significantly more difficult to solve due to the capacity constraints and the combinatorial nature of a MINLP (mixed-integer nonlinear programming) network problem. Moreover, it differs from the SC-COQ model in two main aspects: (1) several business entities can be selected at each echelon of the SC; (2) the components from all selected suppliers enter a plant and are mixed; thus, a shipment to a retailer contains products with components from different suppliers. Therefore, a pooled fraction defective from the selected suppliers is computed. Specialized solution procedures were developed to address this problem. The decision variables are to select the best combination of one or more suppliers, decide which plants of a given set to open, and select the best combination of one or more retailers in order to maximize the total profit and satisfy a minimum quality level for the final product as illustrated in Figure 1.

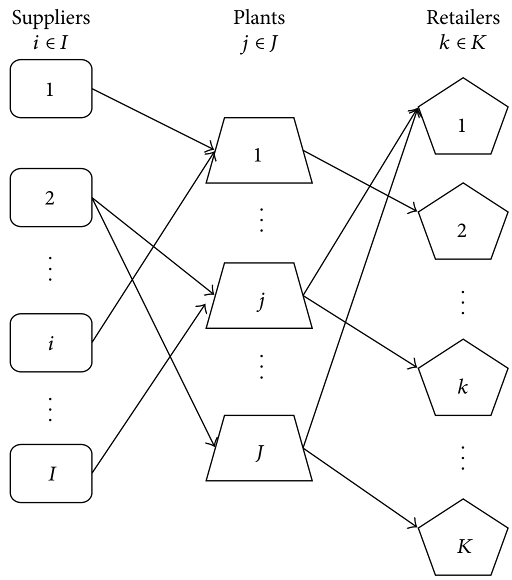



Figure 1: Illustration of a supply chain network.

## 2. The Capacitated SCND-COQ Model

The model assumptions are as follows.

(1) A consumer goods SC, consisting of three echelons: suppliers, manufacturers, and retailers, is modeled.

(2) A single product is modeled.

(3) The objective is to maximize profit.

(4) The overall quality level, QL, (see xvi) is sufficient to represent the quality of the final product.

(5) A new SC is being designed. Modifications to the model would be necessary if implemented for a manufacturing process which is currently working at a specific fraction defective at manufacturing and with an established inspection system.

(6) External failure costs can be tolerable for the firm. This implies direct applicability to products where external failure is not catastrophic for customers. However the overall quality level (QL) can be adjusted to address problems in which external failures are not desirable (e.g., aerospace and pharmaceutical industries).

(7) Suppliers and retailers are external to the manufacturing plant.

(8) It is assumed that a 100% inspection at the end of the manufacturing process is performed to check component conformance. Two types of errors may arise; Type I error involves classifying a good item as defective and Type II error involves labeling a defective item as good. Type I error is not considered in this model because it is not detrimental to customer satisfaction. Type I and Type II errors are defined in this paper as in [9], in terms of inspector fallibility.

(9) All defective products are returned by customers and incur external failure costs.

(10) The relevant operational costs are production, procurement, transportation from supplier to manufacturing plant, transportation from manufacturing plant to retailer, and a fixed cost for opening the plant.

(11) Customer demand at each retailer ($Dem_k$) is known for the study period; demand is determined by the retailers and communicated to the company. Therefore, retailers' capacity is not considered.

(12) Suppliers and manufacturing plants have finite capacity.

(13) At least one supplier, one plant, and one retailer must be selected (simplest supply chain network).

*2.1. Computing the Quality Costs for the Capacitated Model.* An example of the representation of the flow of items through

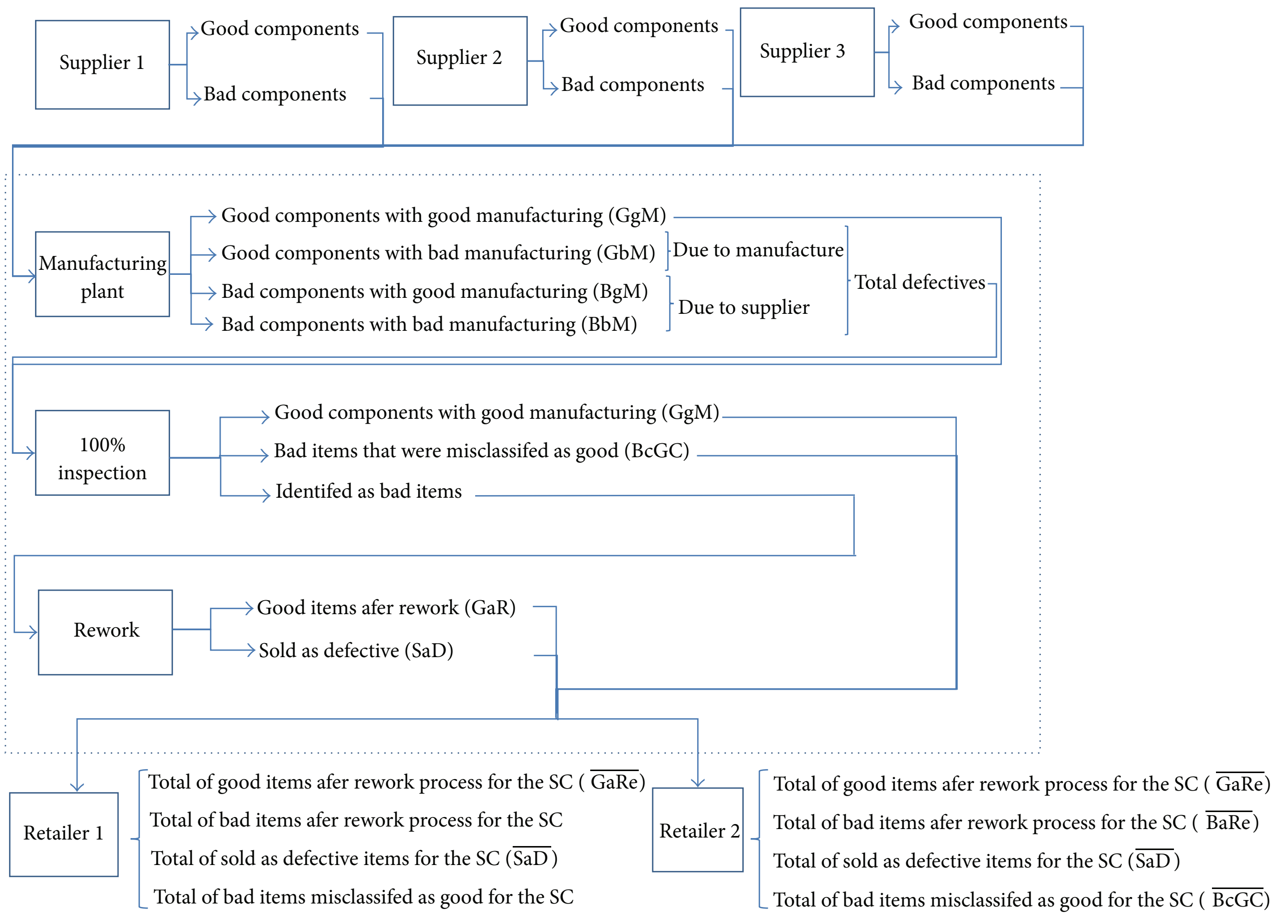



Figure 2: Representation of the flow of items through a supply chain network.

a SC network is shown in Figure 2 where a fixed network of three suppliers, one manufacturing plant, and two retailers is assumed.

The mathematical notation is presented below.

*Sets*

$I$: set of suppliers ($i \in I$).

$J$: set of manufacturing plants ($j \in J$).

$K$: set of retailers ($k \in K$).

*Parameters*

$Dem_k$: captured customer demand for retailer $k \in K$.

$Cap_i$: maximum capacity at supplier $i \in I$ for procuring components.

$Cap_j$: maximum capacity at manufacturing plant $j \in J$ for the production of items.

$Y_{si}$: fraction defective at supplier $i \in I$.

$\overline{Y}s_j = \sum_i Ys_i w_{ij} / \sum_i w_{ij}$: pooled fraction defective of all suppliers shipping products to manufacturing plant $j \in J$.

$Y_{rk}$: fraction defective at retailer $k \in K$.

$p_{jk}$: price per product sold by manufacturing plant $j \in J$ to retailer $k \in K$.

$Pc_{ij}$: direct cost of components shipped from supplier $i \in I$ to plant $j \in J$.

$Po_{ij}$: production cost (base cost) for component from supplier $i \in I$ transformed at manufacturing plant $j \in J$.

$u_{ij}$: cost of transporting one component from supplier $i \in I$ to plant $j \in J$.

$l_{jk}$: cost of transporting one item from plant $j \in J$ to retailer $k \in K$.

$F_j$: fixed cost for operating manufacturing plant $j \in J$.

*Parameters for the COQ Function*

$A_{fj}$: fixed cost for prevention activities at manufacturing plant $j \in J$.

$A_{vi}$: variable cost for prevention activities implemented by supplier $i \in I$.

$A_{vj}$: variable cost for prevention activities implemented by plant $j \in J$.

$Av_{ij}$: variable cost for combined prevention activities at supplier $i \in I$ and manufacturing plant $j \in J$.

$Bf_j$: fixed cost of inspection at manufacturing plant $j \in J$.

$Bv_j$: variable cost of inspection at manufacturing plant $j \in J$.

$Cf_j$: fixed cost for internal failure cost at manufacturing plant $j \in J$.

$Cs_j$: loss incurred due to failure of components procured from supplier to meet quality requirements (replacement costs and payroll costs incurred) at manufacturing plant $j \in J$.

$Cr_j$: rework cost per defective item at manufacturing plant $j \in J$.

$\overline{C}_j$: cost per defective item associated with repair or replacement of the product at manufacturing plant $j \in J$.

$\phi_j$: rework rate at manufacturing plant $j \in J$.

$\bar{l} = (Cost/100)/(Ub - Lb)$: loss coefficient for the Taguchi loss function associated with the cost of working at the specification limit (for the whole network) and the width of the specification.

$P^*_{jk}$: price per "sold as defective" item sold by manufacturing plant $j$ to retailer $k$.

*Decision Variables*

$yI_j$: inspection error rate at the output of manufacturing plant $j \in J$ (continuous variable between 0 and 1).

$yp_j$: fraction defective at manufacturing plant $j \in J$ (continuous variable between 0 and 1).

$Z_i$: binary variable which equals 1 if supplier $i \in I$ is selected, zero otherwise.

$R_k$: binary variable which equals 1 if retailer $k \in K$ is selected, zero otherwise.

$P_j$: binary variable which equals 1 if plant $j \in J$ is open, zero otherwise.

$w^{sp}_{ij}$: number of components shipped from supplier $i \in I$ to manufacturing plant $j \in J$.

$w^{pr}_{jk}$: number of components shipped from manufacturing plant $j \in J$ to retailer $k \in K$.

*Expressions*

(i) $\mathrm{GgM}(w^{sp}_{ij}, yp_j, Z_i, P_j) = \sum_i \sum_j (1 - Ys_i) w^{sp}_{ij} (1 - yp_j) Z_i P_j$ represents good components with successful manufacturing for the whole network.

(ii) $\mathrm{GbM}(w^{sp}_{ij}, yp_j, Z_i, P_j) = \sum_i \sum_j (1 - Ys_i) w^{sp}_{ij}\, yp_j Z_i P_j$ represents good components with defective manufacture for the whole network. These defective products are due to the manufacturing process.

(iii) $\mathrm{BgM}(w^{sp}_{ij}, yp_j, Z_i, P_j) = \sum_i \sum_j Ys_i w^{sp}_{ij} (1 - yp_j) Z_i P_j$ represents bad components with successful manufacture for the whole network. These defective products are due to the supplier.

(iv) $\mathrm{BbM}(w^{sp}_{ij}, yp_j, Z_i, P_j) = \sum_i \sum_j Ys_i w^{sp}_{ij}\, yp_j Z_i P_j$ represents bad components with defective manufacture for the whole network. These defective products are due to the supplier.

(v) $\mathrm{GaRe}(w^{sp}_{ij}, yp_j, yI_j, Z_i, P_j) = \sum_i \sum_j \phi_j (1 - yI_j) w^{sp}_{ij} [(1 - Ys_i) yp_j + Ys_i] Z_i P_j$ is a function that returns the number of good products after successful rework before leaving the plants.

(vi) $\mathrm{SaD}(w^{sp}_{ij}, yp_j, yI_j, Z_i, P_j) = \sum_i \sum_j (1 - \phi_j)(1 - yI_j) w^{sp}_{ij} [(1 - Ys_i) yp_j + Ys_i] Z_i P_j$ is a function that returns the number of defective products which will be sold at a reduced price before leaving the plants.

(vii) $\mathrm{BcGC}(w^{sp}_{ij}, yp_j, yI_j, Z_i, P_j) = \sum_i \sum_j yI_j w^{sp}_{ij} [(1 - Ys_i) yp_j + Ys_i] Z_i P_j$ is a function that returns the bad items that were misclassified as good at inspection before leaving the plants.

(viii) $\overline{\mathrm{SaD}}(w^{pr}_{jk}, yp_j, yI_j, P_j, R_k) = \sum_j \sum_k (1 - \phi_j)(1 - yI_j) w^{pr}_{jk} [(1 - \overline{Y}s_j) yp_j + \overline{Y}s_j] P_j R_k$ is a function that returns the number of defective products which will be sold at a reduced price at retailers. The computation utilizes the pooled fraction defective at suppliers because the items shipped to the retailer may come from different suppliers.

(ix) $\overline{\mathrm{SaD}}_j(w^{pr}_{jk}, yp_j, yI_j, R_k) = \sum_k (1 - \phi_j)(1 - yI_j) w^{pr}_{jk} [(1 - \overline{Y}s_j) yp_j + \overline{Y}s_j] R_k$ is a function that quantifies the number of defective products which will be sold at a reduced price at retailers *per manufacturing plant*.

(x) $\overline{\mathrm{BcGC}}(w^{pr}_{jk}, yp_j, yI_j, P_j, R_k) = \sum_j \sum_k yI_j w^{pr}_{jk} [(1 - \overline{Y}s_j) yp_j + \overline{Y}s_j] P_j R_k$ is a function that returns the number of bad products after the manufacturing process. The computation utilizes the pooled fraction defective at suppliers.

(xi) $\overline{\mathrm{BcGC}}_j(w^{pr}_{jk}, yp_j, yI_j, R_k) = \sum_k yI_j w^{pr}_{jk} [(1 - \overline{Y}s_j) yp_j + \overline{Y}s_j] R_k$ is a function that quantifies the number of bad products after the manufacturing process *per manufacturing plant*.

(xii) $\overline{\mathrm{GaRe}}\,(w^{pr}_{jk}, yp_j, yI_j, P_j, R_k) = \sum_j \sum_k (1 - Yr_k) [(1 - \overline{Y}s_j) w^{pr}_{jk} (1 - yp_j) + \phi_j (1 - yI_j) w^{pr}_{jk} ((1 - \overline{Y}s_j) yp_j + \overline{Y}s_j)]\, P_j R_k$ is a function that returns the number of good products delivered to the customers for the whole network.

(xiii) $\overline{\mathrm{BaRe}}(w^{pr}_{jk}, yp_j, yI_j, P_j, R_k) = \sum_j \sum_k Yr_k w^{pr}_{jk} [(1\text{-}\overline{Y}s_j)(1\text{-}yp_j) + \phi_j (1 - yI_j)((1 - \overline{Y}s_j) yp_j + \overline{Y}s_j)] P_j R_k$ is a function that returns the number of bad products delivered to the customers for the whole network.

(xiv) $\overline{\mathrm{BaRe}}_j(w^{pr}_{jk}, yp_j, yI_j, R_k) = \sum_k Yr_k w^{pr}_{jk} [(1 - \overline{Y}s_j)(1 - yp_j) + \phi_j (1 - yI_j)((1 - \overline{Y}s_j) yp_j + \overline{Y}s_j)] R_k$ is a function that quantifies the number of bad products delivered to the customers for the whole network *per manufacturing plant*.

(xv) $y(w_{jk}^{pr}, yp_j, yI_j, P_j, R_k) = \overline{\text{BaRe}} + \overline{\text{BcGC}} + \overline{\text{SaD}} / \sum_j \sum_k w_{jk}^{pr} P_j R_k$ is the overall fraction defective for the whole supply chain network.

(xvi) $QL_k(w_{jk}^{pr}, yp_j, yI_j, P_j, R_k) = \sum_j (1 - Yr_k)[(1 - \overline{Y}s_j) w_{jk}^{pr}(1 - yp_j) + \phi_j(1 - yI_j) w_{jk}^{pr}((1 - \overline{Y}s_j) yp_j + \overline{Y}s_j)] P_j / \sum_j w_{jk}^{pr} P_j R_k$ is the overall quality level achieved at retailer $k$.

(xvii) $\text{QL}(w_{jk}^{pr}, yp_j, yI_j, P_j, R_k) = \overline{\text{GaRe}} / \sum_j \sum_k w_{jk}^{pr} P_j R_k$ is the overall quality level achieved by the supply chain network.

*2.1.1. Prevention Cost.* Prevention cost is linked to the production of good products after the manufacturing process as given by

$$\begin{aligned} C_P & \left(w_{ij}^{sp}, yp_j, Z_i, P_j\right) \\ &= \sum_j Af_j P_j \\ &\quad + \sum_i \sum_j Av_i \left(Ys_i\right) \left[w_{ij}^{sp} \left(1 - Ys_i\right) Z_i P_j\right] \\ &\quad + \sum_i \sum_j Av_j \left(yp_j\right) \left[w_{ij}^{sp} \left(1 - Ys_i\right) \left(1 - yp_j\right)\right] Z_i P_j \\ &\quad + \sum_i \sum_j Av_{ij} \left(Ys_i, yp_j\right) \\ &\qquad \times \left[w_{ij}^{sp} \left(1 - Ys_i\right) \left(1 - yp_j\right)\right] Z_i P_j, \end{aligned} \tag{1}$$

where $Af_j$ is a fixed cost. The variable cost for prevention activities is divided into three scenarios: (1) the prevention activity is carried out only at suppliers; (2) the prevention activity is implemented only at the manufacturing plants; and (3) the prevention activity is a coordinated action between a supplier and plant. In the first case, $Av_i(Ys_i)$ is a function of the fraction defective at a supplier that returns the cost per unit of good components for prevention activities. Even though this prevention activity is carried out at the supplier, the cost of this prevention activity is incurred by the manufacturing plant. For the second (resp., third) case, $Av_j(yp_j)(Av_{ij}(Ys_i, yp_j))$ is a function of the fraction defective at the selected plant (resp., supplier) that quantifies the cost per unit of good product after the manufacturing process, $\text{GgM}(w_{ij}^{sp}, yp_j, Z_i, P_j)$. These three functions typically increase as the fraction defective at supplier and/or plant decreases (i.e., when the quality level of the supplier and/or manufacturing process improves).

*2.1.2. Appraisal Cost.* A 100% inspection is performed at the end of the manufacturing process to verify conformance. The appraisal costs ($C_A$) are modeled by a fixed cost and a variable cost per item that is classified accurately. Thus, the appraisal cost increases when inspection is more accurate. The appraisal cost is given by

$$\begin{aligned} C_A \left(w_{ij}^{sp}, yI_j, Z_i, P_j\right) &= \sum_j Bf_j P_j \\ &\quad + \sum_i \sum_j Bv_j w_{ij}^{sp} \left(1 - yI_j\right) Z_i P_j, \end{aligned} \tag{2}$$

where $Bf_j$ is a fixed cost and $Bv_j$ is a variable cost. All items going through the system are inspected; thus, $w_{ij}^{sp}$ represents the number of items to be inspected.

*2.1.3. Internal Failure Cost.* The internal failure cost for the capacitated model is given by (3). The first term is a fixed cost ($Cf_j$) for corrective activities. The second term is the internal failure cost due to unsuccessful manufacture computed as rework cost per item ($Cr_j$) times the identified good components with unsuccessful manufacture. It is assumed that these components can be recovered. The third term is the purchasing failure cost computed as the sum of losses incurred due to failure of purchased components to meet quality requirements ($Cs_j$) and rework cost per item ($Cr_j$), multiplied by the number of items identified as defective due to bad components, $\phi_j(1 - yI_j)$ BgM $(w_{ij}^{sp}, yp_j, Z_i, P_j)$, as well as the items identified as defective because of bad components and unsuccessful manufacture, $\phi_j(1 - yI_j)$ BbM $(w_{ij}^{sp}, yp_j, Z_i, P_j)$. The fourth term represents the difference between what would have been the income from sales of nondefective or good products ($p_{jk}$) and the actual income from sales of defective items ($P_{jk}^*$) [10], times the items sold as defective, $\overline{\text{SaD}}(w_{jk}^{pr}, yp_j, yI_j, P_j, R_k)$:

$$\begin{aligned} C_{IF} & \left(w_{ij}^{sp}, w_{jk}^{pr}, yI_j, yp_j, Z_i, P_j, R_k\right) \\ &= \sum_j Cf_j P_j \\ &\quad + \sum_i \sum_j Cr_j \phi_j \left(1 - yI_j\right) \left(1 - Ys_i\right) w_{ij}^{sp} yp_j Z_i P_j \\ &\quad + \sum_i \sum_j \left(Cs_j + Cr_j\right) \phi_j \left(1 - yI_j\right) w_{ij}^{sp} Ys_i Z_i P_j \\ &\quad + \sum_j \sum_k \left(p_{jk} - P_{jk}^*\right) \left(1 - \phi_j\right) w_{jk}^{pr} \left(1 - yI_j\right) \\ &\qquad \times \left[\left(1 - \overline{Y}s_j\right) yp_j + \overline{Y}s_j\right] P_j R_k. \end{aligned} \tag{3}$$

*2.1.4. External Failure Cost.* The external failure costs and opportunity costs are given by

$$\begin{aligned} C_{EF} & \left(w_{jk}^{pr}, yp_j, yI_j, P_j, R_k\right) \\ &= \sum_j \sum_k \overline{C}_j Yr_k w_{jk}^{pr} \left\{\left(1 - \overline{Y}s_j\right) \left(1 - yp_j\right)\right. \end{aligned}$$

$$+\phi_j\left(1-yI_j\right)\times\left[\left(1-\overline{Y}s_j\right)yp_j+\overline{Y}s_j\right]\Big\}P_jR_k$$
$$+\sum_j\sum_k\overline{C}_jyI_jw_{jk}^{pr}\left[\left(1-\overline{Y}s_j\right)yp_j+\overline{Y}s_j\right]P_jR_k$$
$$+\sum_j\bar{l}_j\left(y_j-Lb_j\right)^2P_j, \tag{4}$$

where the first and second terms represent the costs generated by defective items returned by customers, which is the product of $\overline{C}_j$ and the sum of the bad items due to the retailer, $\overline{\mathrm{BaRe}}(w_{jk}^{pr}, yp_j, yI_j, P_j, R_k)$, and the bad items that were classified as good, $\overline{\mathrm{BcGC}}(w_{jk}^{pr}, yp_j, yI_j, P_j, R_k)$. The third term is based on the Taguchi loss function. The loss constant coefficient, $\bar{l}$, depends on the cost for working at the specification limits and the width of the specification [11]. The cost for working at the specification limits ($cp_j$) is computed as a proportion (given by the parameter cost) of the income of the items sold by each manufacturing plant as given by

$$cp_j = Cost\sum_k w_{jk}^{pr}p_{jk} \quad \forall j\in J. \tag{5}$$

The width of the specification is defined by an upper limit ($Ub_j$) which is set to 100 to indicate the allowable deviation from target value. Notice that $y = 100\%$ is the worst case, that is, when the process has 100% defective products. Perfect inspection and manufacturing are assumed in (8) in order to obtain a lower bound or target value ($Lb_j$) for the Taguchi function as given by (6). Thus, the width of the specification is $Ub_j - Lb_j$:

$$Lb_j = \frac{\sum_k Yr_kw_{jk}^{pr}\left[\left(1-\overline{Y}s_j\right)+\phi_j\overline{Y}s_j\right]R_k}{\sum_k w_{jk}^{pr}R_k}\times 100\% \quad \forall j\in J. \tag{6}$$

The loss coefficient is given by

$$\bar{l}_j = \frac{cp_j}{(100)\left(Ub_j-Lb_j\right)} \quad \forall j\in J. \tag{7}$$

The quality characteristic, $y_j$, is the overall percentage defective for plant $j$ as shown in (8). The quality characteristic includes $\overline{\mathrm{BaRe}}_j(w_{jk}^{pr}, yp_j, yI_j, R_k)$, $\overline{\mathrm{BcGC}}_j(w_{jk}^{pr}, yp_j, yI_j, R_k)$, and $\overline{\mathrm{SaD}}_j(w_{jk}^{pr}, yp_j, yI_j, R_k)$. Consider

$$y_j = \left(\overline{\mathrm{BaRe}}_j\left(w_{jk}^{pr}, yp_j, yI_j, R_k\right)+\overline{\mathrm{BcGC}}_j\left(w_{jk}^{pr}, yp_j, yI_j, R_k\right)+\overline{\mathrm{SaD}}_j\left(w_{jk}^{pr}, yp_j, yI_j, R_k\right)\right)\times\left(\sum_j\sum_k w_{jk}^{pr}P_jR_k\right)^{-1}\times 100\% \quad \forall j\in J. \tag{8}$$

In order to compute the opportunity loss for the SC, a relative value of the quality characteristic is obtained by subtracting the target value (a lower bound) from the current overall percentage defective as shown in the last term of (4). In summary, the Taguchi loss is computed for each plant and these costs are summed to obtain the total opportunity loss for the network.

The total COQ is computed as the sum of the prevention, appraisal, and internal and external failure expressions as given by

$$\begin{aligned}\mathrm{COQ}&\left(w_{ij}^{sp}, w_{jk}^{pr}, yI_j, yp_j, Z_i, P_j, R_k\right)\\&= C_P\left(w_{ij}^{sp}, yp_j, Z_i, P_j\right)+C_A\left(w_{ij}^{sp}, yI_j, P_j\right)\\&\quad+C_{IF}\left(w_{ij}^{sp}, yI_j, yp_j, Z_i, P_j, R_k\right)\\&\quad+C_{EF}\left(w_{jk}^{pr}, yp_j, yI_j, P_j, R_k\right).\end{aligned} \tag{9}$$

*2.2. Mathematical Formulation of the SCND-COQ Model.* The objective function is to maximize profit

$$\begin{aligned}&\sum_{j\in J}\sum_{k\in K}w_{jk}^{pr}p_{jk}P_jR_k-\mathrm{COQ}\left(w_{ij}^{sp}, w_{jk}^{pr}, yI_j, yp_j, Z_i, P_j, R_k\right)\\&-\sum_{i\in I}\sum_{j\in J}w_{ij}^{sp}Pc_{ij}Z_iP_j-\sum_{i\in I}\sum_{j\in J}w_{ij}^{sp}Po_{ij}Z_iP_j\\&-\sum_{i\in I}\sum_{j\in J}w_{ij}^{sp}u_{ij}Z_iP_j-\sum_{j\in J}\sum_{k\in K}w_{jk}^{pr}l_{jk}P_jR_k-\sum_{j\in J}F_jP_j\end{aligned} \tag{10}$$

subject to

$$\sum_j w_{jk}^{pr} \le Dem_kR_k; \quad \forall k\in K, \tag{11}$$

$$\sum_i w_{ij}^{sp} = \sum_k w_{jk}^{pr}; \quad \forall j\in J, \tag{12}$$

$$\sum_i w_{ij}^{sp} \le Cap_jP_j; \quad \forall j\in J, \tag{13}$$

$$\sum_j w_{ij}^{sp} \le Cap_iZ_i; \quad \forall i\in I, \tag{14}$$

$$QL_k \ge lR_k; \quad \forall k\in K, \tag{15}$$

$$0\le yI_j\le 1; \quad \forall j\in J, \tag{16}$$

$$0\le yp_j\le 1; \quad \forall j\in J, \tag{17}$$

$$Z_i\in\{0,1\}, \quad P_j\in\{0,1\}, \quad R_k\in\{0,1\}; \quad \forall i\in I, \quad \forall j\in J, \quad \forall k\in K. \tag{18}$$

The objective function given in (10) represents the profit; it is a nonconvex and nonlinear function and has seven components. The first term is the sales revenue. The second term represents the total COQ for the network. The third term represents the direct cost of acquiring components from the selected supplier(s) by the opened manufacturing plant(s). The fourth term represents the operational cost for the components from selected supplier(s) processed at the opened plant(s). The fifth term gives the transportation cost from the supplier(s) to opened plant(s). The sixth term represents the transportation costs from the opened plant(s) to the retailer(s). Lastly, the seventh component determines the fixed cost for opening plants.

Constraints in (11) enforce that demand at retailers is not exceeded. Constraints in (12) ensure that the number of components shipped from suppliers to manufacturing plantsequals the number of items shipped from manufacturing plants to retailers. Constraints in (13) ensure that the plant capacity (in units) is not exceeded. Since the same number of components received from the suppliers is transformed by the manufacturing plants and shipped to the retailers as either good items or sold as defective items, the capacity corresponds to an entry-exit capacity (maximum flow allowed within the plants). Constraints in (14) enforce that the exit capacity (in units) at the suppliers is not exceeded. The exit capacity is the amount of components shipped to manufacturing plants. Constraints in (15) are the quality level constraints; thus, the quality of the final product delivered at each retailer must meet the minimum required quality level; this set of constraints is nonlinear. Constraints shown in (16) and (17) define feasible ranges and binary requirements for the model variables.

## 3. Solution Procedures

The capacitated SCND-COQ model is a constrained mixed-integer nonlinear programming problem (MINLP) which is challenging to solve because it combines all the difficulties of both of its subcategories: the combinatorial nature of mixed integer programming (MIP) and the difficulty of solving non-convex nonlinear problems (NLP). These two subcategories are known as NPO-complete problems [12]; thus, solving MINLP problems can be a daunting task. Metaheuristics have proven to be computationally more efficient than gradient-based nonlinear programming methods for MINLP. Hence, developing an effective metaheuristic-based algorithm to deal with problems of practical and realistic size is preferred.

Five procedures for solving the capacitated SCND-COQ model are described and compared. Three heuristic procedures are based on the serial model and they can be divided into two stages. Stage I consists of finding serial logistic routes, a serial route being a combination of three entities (supplier-plant-retailer), with the highest profit per unit sold, which are to be added to the network at each iteration of the procedure. Once a feasible network is constructed, Stage II consists of evaluating the feasible network configuration using the capacitated SCND-COQ model. The difference among the serial-based procedures lies in how the search of the serial routes to be added to the network at each iteration is performed. The first heuristic procedure enumerates all the serial routes and applies a value-restricted selection for finding the serial logistic route to be added to the network at each iteration of the algorithm. The second procedure uses the local search metaheuristic simulated annealing (SA) for finding the serial logistic route to be added to the network at each iteration of the algorithm. The third procedure uses a population-based metaheuristic, the genetic algorithm (GA), for finding the serial logistic route to be added to the network at each iteration of the algorithm. In addition, two solution procedures which are not based on finding serial logistic routes were developed: (1) an exhaustive enumeration of all possible networks with calls to a GlobalSearch (GS) algorithm and (2) an exhaustive enumeration of all possible networks with calls to a MultiStart (MS) algorithm. A detailed description of each of the five solution procedures is presented in Sections 3.1 and 3.2.

*3.1. Heuristic Procedures Based on Adding Serial Logistic Routes.* The number of possible serial logistic routes is computed as $|I| \times |J| \times |K|$. For instance, the number of serial routes for a problem with 5 suppliers, 3 plants, and 5 retailers is 75, which is considerably less than the number of possible network configurations, 6,727. The total number of possible network configurations is given by the following expression:

$$\sum_{i=1}^{|Z|} C(|Z|, i) \sum_{j=1}^{|P|} C(|P|, j) \sum_{k=1}^{|R|} C(|R|, k), \qquad (19)$$

where $C(n, r)$ is the number of ways in which $r$ items can be selected from among $n$ items without replacement.

The heuristic procedures based on the serial model rely on the following idea: a network can be constructed by (1) choosing the serial logistic route with the highest profit per unit sold when sending the maximum possible amount of items through that route, (2) adding that serial route to the network, (3) updating the remaining capacities, and repeating the process.

The heuristic procedures serve to construct a feasible network and to determine the amount of items to be sent (Stage I). Despite the ease of implementation and speed of these heuristic procedures for constructing a network, some limitations exist. For each serial route, the internal decision variables error rate at inspection ($yI_j$) and fraction defective at manufacturing ($yp_j$) must be determined. This implies that, for example, a manufacturing plant included in multiple serial routes could have different values of the internal decision variables. Operating the real system in this way may not be feasible or desirable. To remedy this, a reoptimization of the internal decision variables ($yI_j$ and $yp_j$) is performed for the constructed feasible SC network by using the capacitated SCND-COQ model (Stage II). The profit achieved by this network is taken as the best-found solution for the capacitated model. The general procedure for the heuristic serial-based solution methods is depicted as follows.

*Stage I*

(1.1) Create a list of all possible serial routes ($|I| \times |J| \times |K|$).

(1.2) Compute the quality level attained by each serial route, eliminate the routes that do not meet the minimum level in (15), and save the resultant matrix with all the feasible serial routes (**PS** matrix).

(1.3) Determine the maximum flow that can be sent through a route by evaluating the following: $\min\{\sum_{i\in I} Cap_i, \sum_{j\in J} Cap_j, \sum_{k\in K} Dem_k\}$.

(1.4) Prelocate the vector with not opened plants (NOP). Since the same plant can be selected in several serial routes (as long as the remaining plant's capacity is greater than zero), this vector avoids taking the fixed cost for opening a plant into account more than once.

(1.5) The search for additional serial routes to be added to the network continues until one of the five following cases occurs: nonpositive profit is obtained, the sum of the capacities of the suppliers is exhausted, the sum of the capacities of the plants is exhausted, the demand is satisfied, or there are no more feasible remaining routes to select from (the updated **PS** matrix is empty).

(1.5.1) The search is performed by using one of the following procedures: (1) evaluation of all the serial routes and a greedy or value-restricted constructive (VRC) procedure, (2) simulated annealing (SA), and (3) the genetic algorithm (GA). The details of the procedures are described in Sections 3.1.1, 3.1.2, and 3.1.3.

(1.5.2) Update the remaining capacities and demands.

(1.5.3) One or more of these three cases may occur: one supplier is saturated, one plant is saturated, or the demand at one retailer is fully satisfied. In each case, the business entities that were saturated are eliminated from the set of potential business entities and all the routes that include these business entities are eliminated from the matrix with possible serial routes (**PS** matrix).

(1.5.4) Update the NOP vector each time a plant is selected. For instance, if the selected route contains a plant that was already opened in a previous iteration, then the additional fixed cost is zero; otherwise, if the plant is in the NOP vector, then a fixed cost is incurred for opening that plant.

(1.6) Store results.

*Stage II*

The network with flows formed by adding serial routes is evaluated by using the capacitated SCND-COQ model and the internal continuous variables are reoptimized. It is worth noting that the network and flows found in Stage I are not modified.

(2.1) Reoptimize the internal continuous variables associated with the opened manufacturing plants by using the GlobalSearch algorithm in MATLAB.

*3.1.1. Value-Restricted Constructive Procedure for Selecting Serial Routes.* For each serial route (rows in the possible serial routes matrix, that is, the **PS** matrix), the internal decision variables ($yI_j$ and $yp_j$) that minimize the total COQ are obtained by using a nonlinear solver, FMINCON with the interior-point algorithm of MATLAB. The total profit and the profit per unit sold are computed for each serial route. The profit per unit sold is used to select a serial route; this avoids selecting the route that generates the maximum profit based on volume. Ties are broken by selecting the route that yields a higher total profit.

Value-restricted selection (VRS) was used to choose the serial route to be added to the network at each iteration of the procedure. The selection of the logistic route at each iteration is determined by choosing a random route from a restricted candidate list (RCL). The RCL was used as a way to avoid a greedy choice as in the GRASP (greedy randomized adaptive search procedure) metaheuristic; refer to [13–15].

The RCL contains the routes with the higher values of profit per unit sold. Let $\delta$ be a real value such that $0 \le \delta \le 1$. The RCL consists of all the routes $e$ such that the greedy function $c(e)$, total unit profit, is $c(e) \ge c^* - \delta(c^* - c_*)$. When $\delta = 0$, then the selection is greedy and when $\delta = 1$, the selection is completely randomized. The serial-based heuristic procedure using a VRS for selecting the serial routes at each iteration and the GlobalSearch method for optimizing the internal decision variables of the constructed network is named SVRC (serial value-restricted constructive procedure) and was implemented using MATLAB. The SVRC algorithm with $\delta = 0.1$ ($\delta$ was obtained from previous computational experiments) is named SVRC1 and with $\delta = 0$ (which is a greedy procedure) is named SVRC2).

*3.1.2. SA for Selecting Serial Routes.* Descriptions of the general SA procedure can be found in [16–18]. The SA-based solution procedure used to find the serial route that maximizes unit profit while satisfying a required quality level of the final product is described in [8]. The SA-based procedure chooses the serial route that yields the highest-found profit per unit sold at each iteration. In contrast to the SVRC procedures, SA search procedure does not enumerate all the serial routes at each iteration of the heuristic procedure. The SA serial-based heuristic procedure using the GlobalSearch method for optimizing the internal decision variables of the constructed network is named SSA1 (serial simulated annealing method version 1) and was implemented using MATLAB.

Two additional variations of this method were developed: SSA2 and SSA3. SSA2 has as decision variables the selection of serial routes while the internal continuous variables ($yI_j$ and $yp_j$) are optimized using the nonlinear solver FMINCON with the interior-point algorithm of MATLAB. The difference between SSA1 and SSA2 is the use of FMINCON for the optimization of the continuous variables.

SSA3 also has as decision variables the selection of serial routes while the internal continuous variables ($yI_j$ and $yp_j$) are optimized using the nonlinear solver FMINCON with the interior-point algorithm of MATLAB. The difference between

SSA1 and SSA2 versus SSA3 lies in the state representation. SSA1 and SSA2 determine the next neighborhood solution by indicating how many indexes apart from the current index of the **PS** matrix (matrix where rows represent feasible combinations of entities) this next solution could be. For SSA3, the state is represented by a vector containing three indexes, and each index represents an entity and its value ranges from 1 to the amount of entities of each type. When an unfeasible route is selected (that is not in **PS** matrix), a penalty is performed. In such cases, the value of the objective function is set to zero. This state representation for determining the next neighboring solution was implemented in SSA3 to gain insight about the impact of the state representation on the solution procedure's performance.

*3.1.3. GA for Selecting Serial Routes.* Genetic algorithms, introduced by Holland [19], refer to a class of adaptive search procedures based on the principles derived from natural evolution and genetics. A GA solution procedure to find the serial route that maximizes unit profit and satisfies a required quality level of the final product is described in [8]. The GA based procedure chooses the serial route that yields the highest profit per unit sold at each iteration. In contrast to the SVRC procedures, the GA search procedure does not enumerate all the serial routes at each iteration of the heuristic procedure. The serial-based heuristic procedure using GA and GlobalSearch is named SGA1 (serial genetic algorithm method version 1) and was implemented using MATLAB. Similar to SSA, SGA has three variations. SGA2 and SGA3 algorithms call the nonlinear solver FMINCON with the interior-point algorithm of MATLAB for the optimization of the internal continuous variables while SGA1 optimizes all decision variables (selection of entities and internal continuous variables). For SGA1 and SGA2, the representation of the individual genotype consists of binary numbers that in decimal base represent an index in the **PS** matrix. In SGA3, the individual genotype is segmented in three parts that define the serial route; each part consists of binary numbers that in decimal base represent an independent entity (supplier, plant, or retailer). Equivalently to SSA3, this representation was implemented to gain insight about the impact of the individual representation on the solution procedure's performance. During the optimization process (selection of serial routes), evolutionary operations (selection, crossover, and mutation) are performed over the population in order to improve the population fitness over generations. A binary tournament selection and a one-point crossover method were adopted in the present work for all GA-based solution procedures. When an unfeasible route is selected (that is not in **PS** matrix), a penalty is performed. In such cases, the value of the objective function is set to zero.

*3.2. Exhaustive Enumeration Procedures.* Exhaustive enumeration consists of listing all possible SC networks. For each SC network, the MATLAB GlobalSearch (GS) and the MultiStart (MS) algorithms are used to solve for the amount of items to be sent between suppliers and manufacturing plants and between plants and retailers (i.e., $w_{ij}^{sp}$ and $w_{jk}^{pr}$ matrices, resp.) as well as the variables representing the quality system within the manufacturing plant (i.e., $yI_j$ and $yp_j$). Both algorithms have similar approaches; they run a nonlinear solver (FMINCON) from multiple starting points and try to maximize the total profit of the SC network while satisfying the SCND-COQ model constrains. The solution with the highest profit (related to one of the many SC network configurations with optimized variables) is reported as the best found.

It should be noted that the exhaustive enumeration of all possible SC networks does not necessarily mean that all the possible solutions in the search space are evaluated, as each possible SC network contains infinite possible solutions, which depend on the values of the decision variables ($w_{ij}^{sp}$, $w_{jk}^{pr}$, $yI_j$, and $yp_j$). Moreover, the nonlinear solvers will not always find optimal solution because the solvers may return a local maximum.

Since the number of networks grows exponentially as the number of business entities increases, as shown in (19) and as described in Section 3.1, exhaustive enumeration can only be performed for small problems and heuristic procedures are needed to deal with problems of practical and realistic size. For instance, a problem with 10 suppliers, 15 manufacturing plants, and 2 retailers has $3.5148 \times 10^9$ possible SC network configurations and 237 decision variables.

## 4. Experimental Study

To investigate the effectiveness of the developed solution procedures, a variety of network sizes (as shown in Table 1) were solved. Moreover, three classes of instances were defined and five instances from each class were randomly generated. The pool of test problems consists of 60 problems (15 for each problem size).

In order to compare the performance of the solution procedures against optimal solutions, specially constructed instances (problem class I described in Subsection 4.1.1) were constructed such that the optimal solution is known in advance. Additional class problems where the optimal solution is not known were also generated (classes II and III). A comparison of the solution procedures relative to each other was performed for classes II and III.

*4.1. Test Problem Generation.* The general approach used to generate instances is described next. The data used in test problems was generated randomly from a uniform distribution between the low and high levels documented in Table 2. The minimum required quality level ($l$) is fixed at 0.85 for all test instances. When the required quality level is too low, the solution space will be larger and finding the near-optimal solutions will be more difficult. When the required quality level is too high, the solution space will be smaller and highly constrained. The interested reader can obtain the test problems used in the development of this paper from the authors.

The function of the variable cost for prevention activities was considered as in scenario 3 (as described in Section 2.1.1),

Table 1: Test problems.

| Number of suppliers ($I$) | Number of plants ($P$) | Number of retailers ($R$) | Number of constraints | Number of decision variables |
|---|---|---|---|---|
| 3 | 2 | 3 | 13 | 24 |
| 5 | 4 | 6 | 25 | 67 |
| 8 | 8 | 10 | 44 | 186 |
| 35 | 20 | 35 | 145 | 1530 |

Table 2: Ranges of the parameters used to generate the instances.

| Input parameter | Low level | High level |
|---|---|---|
| Fraction defective at supplier ($Ys_i$) | 0.05 | 0.2 |
| Fraction defective at retailer ($Yr_k$) | 0.05 | 0.1 |
| Extra percentage (extra) in price ($p_{jk}$) | 1.2 | 1.3 |
| Procurement costs ($Pc_{ij}$) | 50 | 120 |
| Production costs ($Po_{ij}$) | 70 | 130 |
| Transportation costs: $u_{ij}$ and $l_{jk}$ | 3 | 12 |
| Fixed cost for opening manufacturing plants ($F_j$) | 80,000 | 120,000 |
| Fixed costs: $Af_j$, $Bf_j$, and $Cf_j$ | 5,000 | 15,000 |
| Rework cost ($Cr_j$). | 70 | 90 |
| Loss incurred owing to failure of purchased components ($Cs_j$) | 0.45 of average $Pc_{ij}$ | 0.55 of average $Pc_{ij}$ |
| Variable cost for prevention activities ($Av_{ij}(Ys_i, yp_j)$) | $Bv_j$ | 5 $Bv_j$ |
| Variable cost for appraisal/inspection activities ($Bv_j$) | 5 | 5 |
| Price per "sold as defective" items ($P^*_{jk}$) | 1/4 $p_{jk}$ | 3/4 $p_{jk}$ |
| Cost for computing Taguchi loss function for the network ($Cost$) | 1/10 | 1/3 |
| Cost per defective item ($\overline{C}_j$) | 1/4 of avg. price | 1/2 of avg. price |
| $Dem_k$, $Cap_i$, and $Cap_j$ | 50,000 | 80,000 |

where $Av_{ij}(Ys_i, yp_j) = c$ ($c$ is a constant) for the test problems. The rework rate ($\phi_j$) is set to $Cr_j/100$. The price ($p_{jk}$) is calculated using (20). The price considers an extra percentage (extra) for revenue and for covering administrative costs as shown in Table 2. The vectors of average cost for all potential suppliers for plant $j$ are differentiated from the matrices by using a bar ($\overline{Av}_j$ for prevention activities; $\overline{Pc}_j$ for procurement costs; and $\overline{u}_j$ and $\bar{l}_j$ for transportation costs) and used in

$$p_{jk} = \text{round}\left(\sum_k \left(\overline{Av}_j + Bv_j + Cr_j + Cs_j + \overline{Pc}_j + Po_j + \overline{u}_j + \bar{l}_j + \frac{\left(Af_j + Bf_j + Cf_j\right)}{Dem_k}\right)\right)\text{extra}. \quad (20)$$

The cost for computing the Taguchi loss function ($Cost$) is given by

$$Cost = \sum_{k\in K} w^{pr}_{jk} p_{jk} U(a,b) \quad \forall j \in J, \quad (21)$$

where $p_{jk}$ is the average price and $U(a,b)$ is the continuous uniform distribution of the low and high levels in Table 2 for $Cost$.

*4.1.1. Class I Test Instances.* In class I problem instances, the optimal solution is a serial route that satisfies all the demand at the selected retailers. This was accomplished by generating instances as described above but setting the extra percentage (extra) to 0.5, instead of the interval shown in Table 2.

Once the parameter values are generated from uniform distributions with ranges as shown in Table 2, some of the parameters associated with the business entities in the optimal serial route ($Z^*$ will denote the optimal supplier selection, $P^*$ the optimal plants, and $R^*$ the optimal retailers) are modified in order to force the optimal solution to be a specific serial route. A ratio $\beta = 0.6$ will be used to adjust the parameters of $Z^*$, $P^*$, and $R^*$ so that the cost parameters of these entities are favorable (significantly lower than the rest of the values of the cost parameters of other business entities) to make this specific serial route the optimal solution. The optimal business entities' cost parameters are modified by taking the low level in the ranges in Table 2 and multiplying by $\beta$. In this way, the costs are $(1-\beta)\%$ less than the rest of the cost parameters. The fraction defective at the supplier and retailer are also decreased by $(1-\beta)\%$ for $Z^*$ and $R^*$.

The rework rate ($\phi_j$) is modified for $P^*$ by considering the maximum rework rate among all the plants and then dividing by 100. The demand at $R^*$ is set to the highest demand generated multiplied by $1+\beta$. The capacities at $Z^*$ and $P^*$ are set such that they exactly match the demand at $R^*$.

The sales price is set to three times the maximum generated price. The price of "sold as defective" items of $P^*$ is computed as the sales price multiplied by the high level of $P^*_{jk}$ in the range shown in Table 2.

The instances created by using the above procedure are verified by enumerating and evaluating all the possible serial routes. The maximum flow is sent through the route and the internal continuous variables are found by using a nonlinear solver to obtain the profit. The solution with the maximum profit should select $Z^*$, $P^*$, and $R^*$; otherwise, the instance is not used for testing.

*4.1.2. Class II Test Instances.* For class II problem instances, the opening of all the business entities to satisfy the demand at retailers is expected. This was accomplished by increasing the price of the final items and modifying capacities so that the retailers limit the flow.

The parameter values are randomly generated from uniform distributions with ranges as shown in Table 2. However, the price for class II problems ranges from 1.9 to 2. The higher prices are conducive to networks with positive profit (even when the quality costs are higher in some networks than others). Moreover, the capacities are modified so that retailers limit the amount of items to be sent. The sum of the randomly generated retailer capacities $Dem_k$ is multiplied by 1.1 and divided by the number of suppliers to obtain the capacity at each supplier. This calculation is repeated for plants. Thus, the suppliers and plants have enough capacity to satisfy the demand at retailers.

*4.1.3. Class III Test Instances.* The data for class III problems was generated randomly from a uniform distribution between the low and high levels documented in Table 2 as described in the general approach presented at the beginning of Section 4.1.

*4.2. Computational Results.* Instances were solved using each of the five proposed solution procedures. Since the number of SC network configurations grows exponentially as the size of the problem increases, exhaustive enumeration procedures are used only for the 3 × 2 × 3 test problem size. For the rest of the problems, a comparison of SVRC, SSA, and SGA procedures (including its different versions) is performed.

*4.2.1. Runs and Parameter Setting for the Serial-Based Solution Procedures.* The SA algorithm is run 5 times in order to select one serial route with the best-found objective value at each iteration of the heuristic procedure. The GA algorithm is also run 5 times. The VRC procedure performs an enumeration of all the possible serial routes and selects one from the RCL at each iteration of the heuristic procedure. The whole procedure (steps (1.5) and (1.6) in Section 3.1) is run 5 times and the network configuration with the best accumulated profit was reported.

The heuristic parameter for SVRC1 and SVRC2 is $\delta$. The heuristic parameters for SSA1, SSA2, and SSA3 (phase I) are the neighborhood size for the internal decision variables (*neigsize*), the initial system temperature ($T_o$), the rate of cooling ($\alpha$), the number of accepted trials (*AccepTrials*), the maximum number of trials (*MaxTrials*), and the maximum number of Markov Chains (*MaxChains*). The heuristic parameters for SGA1, SGA2, and SGA3 (phase I) are the number of generations (*gen*), probability of mutation (*pm*), the probability of crossover (*pc*), and the initial population (*pop*) which is computed as a percentage of the **PS** matrix. All heuristic parameters were tuned using statistical designs of experiments. Table 3 shows the values of the heuristic parameters.

In previous studies, the COQ function, based on percentage of defective units, is assumed to be given. This paper deals with the development of an SCND-COQ model that computes the COQ for a whole SC based on internal decisions within the manufacturing plant, such as fraction defective at the manufacturing plant and error rate at inspection. No previous work has addressed how the COQ functions were obtained while taking internal operational decisions within the SC. Moreover, the proposed SCND-COQ model computes quality costs for the whole SC considering the interdependencies among business entities, whereas previous works have assumed independent COQ functions for each node of the SC.

*4.2.2. Comparison among Solution Procedures.* Tables 4–7 show the results for each problem size. Performance was measured by solution quality, number of evaluations of the objective function, and computational time in CPU seconds. Solution quality is characterized in two ways: (a) the average best-found profit (Avg_Profit) over 5 instances obtained by each solution procedure and (b) the average percentage deviation from the optimal solution for class I and from the best-found solution for classes II and III (Avg%dev) over the same 5 instances. The deviation at each instance is computed as [(optimal solution − Avg_Profit)/optimal solution)] × 100% for class I and as [(best-found solution − Avg_Profit)/best-found solution)] × 100% for classes II and III.

For the GS and MS procedures, the best-found profit, achieved from one of all possible SC network configurations, is the one reported. For SVRC1 and SVRC2 (greedy approach), only 1 run is performed in Stage 1 and the profit achieved by the network in Stage 2 is the one reported as the final objective value. For SSA/SGA, at each iteration, the serial logistic route with the best-found profit in 5 runs is the one added to the network in Stage 1; the profit obtained by the network in Stage 2 is the one reported.

The average number of evaluations of the objective function over 5 instances (Avg_Evals) is also considered as a performance measure. For the GS and MS procedures, the number of evaluations of the objective function is determined by the total number of times that the nonlinear solver is called while solving the problem instance. For SVRC, SSA, and SGA, the number of evaluations is computed as the sum of the number of evaluations performed by the heuristic procedures when constructing the network and the number of evaluations performed by the nonlinear solver while reoptimizing the internal continuous variables for the capacitated model.

Table 3: Heuristic parameters.

| Solution procedure | Heuristic parameters | | | | |
|---|---|---|---|---|---|
| SVRC1 | $\delta = 0.1$ | | | | |
| SVRC2 | $\delta = 0$ | | | | |
| SSA1 | neigsize = [0.45 0.45 0.1 size (PS, 1)] | $T_o = 1 \times 10^5$ | $\alpha = 0.8$ | Accep/MaxTrials = 30 | MaxChains = 3 |
| SSA2 | neigsize = [0.1 size (PS, 1)]<br>neigsize == [0.3 size (PS, 1)]<br>for 8 × 8 × 10, classes II and III | $T_o = 1 \times 10^5$ | $\alpha = 0.8$ | Accep/MaxTrials = 30;<br>50 for 8 × 8 × 10 class I;<br>and 80 for classes II and III | MaxChains = 2 |
| SSA3 | neigsize = [2 2 2]<br>neigsize == [3 3 3]<br>for 8 × 8 × 10, classes II and III | $T_o = 1 \times 10^5$ | $\alpha = 0.8$ | Accep/MaxTrials = 30;<br>50 for 8 × 8 × 10 class I;<br>and 80 for classes II and III | MaxChains = 2 |
| SGA1 | pm = 0.01 | pc = 0.95 | gen = 20 | pop = 20% of **PS** | |
| SGA2 | pm = 0.01 | pc = 0.95 | gen = 10 | pop = 10% of **PS** | |
| SGA3 | pm = 0.01 | pc = 0.95 | gen = 10 | pop = 10% of **PS** | |

Note: for the 35 × 20 × 35 size, SGA2 and SGA3 have gen = 25 and pop = 120 individuals. The number of runs was decreased from 5 to 3.

Finally, the average computational time (Avg_Time) is the average processing time duration in CPU seconds that is required for each solution procedure over 5 instances. The computational time considers the entire solution procedures, that is, phase I and phase II. The computer used for the computational experiments was a Sager NP8130 with an Intel i7 2720QM processor operating at 3.3 GHz, with 16 GB of memory DDR3 on an Intel HM65 chipset motherboard.

Table 4 gives computational results for the first problem size. For class I problems, it can be observed that most algorithms reached the optimal solutions in all five instances, with the exceptions being MS and SGA1. The GS (exhaustive enumeration procedure) was able to find the optimal solution for class I instances but was unable to find good solutions for classes II and III, thus, denoting a limitation of such procedure to tackle complex instances.

For class II problems, based on Avg_Profit, the SVRC2, SSA2, SSA3, SGA2, and the SGA3 reach the same average solution (best-found solution). For class III problems, the best-found solution was found by SVRC2, SSA1, SSA2, SSA3, SGA2, and SGA3. GS, MS, and SVRC1 present an Avg%dev greater than 6.6%.

Table 5 shows the results for the 5 × 4 × 6 problem size. As observed previously for class I problems, most algorithms reached the optimal solution, with the exception being SGA1 and SGA2. For class II problems, SSA2 found the best solutions on average, followed by SVRC2 and SGA2 (Avg%dev of 0.19% and 0.29%, resp.). Similarly, for class III, SSA2 found the best solutions on average, followed by SVRC2, SSA1, and SGA3. This shows that suboptimal selections of serial routes in SSA phase I may lead to a more profitable network configuration (as shown in SSA2) than the profit obtained by using a greedy approach (SVRC2).

Table 6 shows the results for the 8 × 8 × 10 problem size. For class I problems, most algorithms reached the optimal solution, with the exception being SSA3 and SGA1. For class II, SSA1 found the best solutions on average, closely followed by SVRC2, SSA2, and SGA3. For class III, SSA1 found the best solutions on average, followed by SGA1. It is noteworthy that for class III (considered difficult), the heuristic procedures (SSA1 and SGA1) outperform the enumeration-greedy approach (SVRC2) in average profit and computational time.

The results for the largest instance (35 × 20 × 35) are shown in Table 7. For class I, it was expected that SVRC2 outperformed the rest of the procedures because the optimal solution is a serial route; however, the SSA1 reached solutions with an Avg%dev of 1.64%. For class II, the best-found profits are obtained by SGA1 and SGA3 closely followed by SSA1 and SVRC2. For class III, SVRC2 provides the best found solutions followed by SGA1 and SGA2. In general, the algorithms that produce best solutions on average for large problems are SVRC2, the three GA-based algorithms, and SSA1. The computational burden required by the SVRC2 and the GA-based procedures is significantly greater than the time required by the SSA1 (2.68 hrs on average per instance for SVRC2 and 2.24 hrs on average for the SA-based procedures versus less than 4 minutes for SSA1). SSA1 is able to reach good solutions with a percentage deviation from the best-found average solution for all classes of 0.48%.

To conclude, GS and MS perform poorly, even for the smallest problem size, and require substantial computational effort. The computational time for SVRC2 is for only one run of the whole procedure and the CPU time for the remaining procedures includes the 5 runs in phase I. For realistically sized problems, the practitioner could use the SSA1 which requires less computational time than the remaining procedures to obtain good quality solutions or the GA-based procedures which require more computational effort but obtain the higher profits on average. Interestingly, the index-independent representation is not adequate for the SA-based procedure (SSA3) but improves the average profits reached by SGA3.

## 5. Conclusions and Future Research

The capacitated SCND-COQ model selects several business entities at each echelon of the SC and allows the modeling

Table 4: Results for problem size $3 \times 2 \times 3$.

| | GS | MS | SVRC1 | SVRC2 | SSA1 | SSA2 | SSA3 | SGA1 | SGA2 | SGA3 |
|---|---|---|---|---|---|---|---|---|---|---|
| Class I | | | | | | | | | | |
| Avg_Profit | **41,806,816.26** | 41,709,103.64 | **41,806,816.26** | **41,806,816.26** | **41,806,816.26** | **41,806,816.26** | **41,806,816.26** | 40,569,521.42 | **41,806,816.26** | **41,806,816.26** |
| Avg%dev | **0.00** | 0.26 | **0.00** | **0.00** | **0.00** | **0.00** | **0.00** | 2.57 | **0.00** | **0.00** |
| Avg_Evals | **14,432,210.20** | 16,986,412.80 | **4,314.80** | **925.68** | **7,013.00** | **4,400.00** | **4,481.00** | 4,404.00 | **3,706.40** | **4,278.00** |
| Avg_Time | **12,266.16** | 13,906.77 | **5.40** | **1.09** | **3.20** | **16.89** | **31.83** | 2.59 | **12.59** | **20.67** |
| Class II | | | | | | | | | | |
| Avg_Profit | 43,977,372.68 | 40,721,911.06 | 44,986,134.32 | **47,920,297.79** | 47,390,930.58 | **47,920,297.79** | **47,920,297.79** | 47,576,092.96 | **47,920,297.79** | **47,920,297.79** |
| Avg%dev | 8.92 | 14.66 | 5.37 | **0.00** | 1.16 | **0.00** | **0.00** | 0.57 | **0.00** | **0.00** |
| Avg_Evals | 17,880,237.60 | 16,848,438.00 | 6,656.60 | **1,711.56** | 19,915.40 | **12,811.20** | **12,864.00** | 13,345.60 | **9,601.00** | **8,850.00** |
| Avg_Time | 12,414.88 | 11,740.01 | 8.13 | **2.05** | 9.48 | **93.64** | **100.98** | 8.02 | **47.20** | **72.90** |
| Class III | | | | | | | | | | |
| Avg_Profit | 9,763,747.96 | 9,488,786.92 | 9,778,795.24 | **10,475,602.91** | **10,475,602.91** | **10,475,602.91** | **10,475,602.91** | 10,221,015.40 | **10,475,602.91** | **10,475,602.91** |
| Avg%dev | 6.63 | 6.62 | 8.02 | **0.00** | **0.00** | **0.00** | **0.00** | 3.74 | **0.00** | **0.00** |
| Avg_Evals | 9,420,031.20 | 19,270,506.00 | 5,155.60 | **915.28** | **10,069.40** | **4,843.40** | **5,240.40** | 7,119.80 | **3,630.80** | **4,195.80** |
| Avg_Time | 6,944.75 | 13,079.87 | 6.82 | **1.42** | **3.75** | **44.34** | **59.39** | 3.92 | **21.77** | **36.30** |

The bold font in the tables refers to the algorithms that reached best solutions.

Table 5: Results for problem size $5 \times 4 \times 6$.

| | SVRC1 | SVRC2 | SSA1 | SSA2 | SSA3 | SGA1 | SGA2 | SGA3 |
|---|---|---|---|---|---|---|---|---|
| Class I | | | | | | | | |
| Avg_Profit | **52,844,809.95** | **52,844,809.95** | **52,844,809.95** | **52,844,809.95** | **52,844,809.95** | 49,623,191.12 | 51,243,573.72 | **52,844,809.95** |
| Avg%dev | **0.00** | **0.00** | **0.00** | **0.00** | **0.00** | 5.58 | 3.08 | **0.00** |
| Avg_Evals | **11,037.40** | **2,008.68** | **10,295.20** | **4,928.20** | **7,593.40** | 8,846.60 | 6,951.40 | **5,729.80** |
| Avg_Time | **26.95** | **5.61** | **5.65** | **46.68** | **44.73** | 5.06 | 21.88 | **17.27** |
| Class II | | | | | | | | |
| Avg_Profit | 118,455,380.46 | 125,070,837.06 | 124,353,125.54 | **125,287,428.28** | 123,051,779.22 | 123,204,151.47 | 124,955,144.86 | 123,927,909.74 |
| Avg%dev | 5.58 | 0.19 | 0.74 | **0.00** | 1.70 | 1.76 | 0.29 | 0.98 |
| Avg_Evals | 29,055.40 | 5,049.64 | 42,555.60 | **24,460.00** | 34,449.40 | 46,363.40 | 18,697.20 | 17,002.40 |
| Avg_Time | 74.04 | 17.07 | 14.70 | **324.78** | 317.41 | 24.52 | 133.09 | 146.33 |
| Class III | | | | | | | | |
| Avg_Profit | 39,465,124.69 | 42,810,167.13 | 42,810,167.13 | **42,894,361.03** | 40,157,925.23 | 40,128,165.86 | 42,796,773.97 | 42,810,167.13 |
| Avg%dev | 8.16 | 0.20 | 0.20 | **0.00** | 6.95 | 6.23 | 0.23 | 0.20 |
| Avg_Evals | 22,111.00 | 5,210.44 | 38,800.00 | **20,993.80** | 17,804.60 | 37,709.60 | 13,984.20 | 13,701.60 |
| Avg_Time | 60.39 | 15.60 | 13.89 | **220.63** | 191.32 | 19.02 | 111.73 | 110.97 |

The bold font in the tables refers to the algorithms that reached best solutions.

Table 6: Results for problem size $8 \times 8 \times 10$.

| | SVRC1 | SVRC2 | SSA1 | SSA2 | SSA3 | SGA1 | SGA2 | SGA3 |
|---|---|---|---|---|---|---|---|---|
| Class I | | | | | | | | |
| Avg_Profit | **55,003,836.49** | **55,003,836.49** | **55,003,836.49** | **55,003,836.49** | 53,822,605.31 | 48,794,936.66 | **55,003,836.49** | **55,003,836.49** |
| Avg%dev | **0.00** | **0.00** | **0.00** | **0.00** | 2.02 | 11.21 | **0.00** | **0.00** |
| Avg_Evals | **42,368.80** | **9,186.64** | **9,424.00** | **8,717.80** | 7,919.80 | 22,529.20 | **9,453.20** | **9,871.60** |
| Avg_Time | **149.10** | **29.28** | **3.82** | **49.93** | 66.15 | 6.65 | **83.40** | **76.31** |
| Class II | | | | | | | | |
| Avg_Profit | 210,780,559.55 | 234,647,930.81 | **234,741,055.90** | 234,647,930.81 | 201,791,597.82 | 234,518,007.14 | 234,524,750.57 | 234,647,930.81 |
| Avg%dev | 10.33 | 0.15 | **0.11** | 0.15 | 14.49 | 0.19 | 0.20 | 0.15 |
| Avg_Evals | 174,373.80 | 32,888.00 | **98,987.40** | 82,037.80 | 54,223.60 | 237,543.60 | 79,051.60 | 83,976.00 |
| Avg_Time | 791.55 | 115.83 | **48.11** | 1,487.61 | 1,027.24 | 107.95 | 1,239.64 | 1,575.92 |
| Class III | | | | | | | | |
| Avg_Profit | 77,799,129.47 | 88,500,880.62 | **89,252,162.29** | 88,444,907.51 | 76,311,740.30 | 88,883,338.00 | 88,297,371.17 | 88,444,907.51 |
| Avg%dev | 13.47 | 0.96 | **0.16** | 1.03 | 15.29 | 0.55 | 1.18 | 1.03 |
| Avg_Evals | 166,321.40 | 35,431.92 | **96,052.80** | 88,296.60 | 63,744.00 | 245,791.60 | 87,427.60 | 87,802.00 |
| Avg_Time | 528.86 | 126.18 | **35.45** | 1,705.61 | 1,197.31 | 102.84 | 1,279.54 | 1,212.89 |

The bold font in the tables refers to the algorithms that reached best solutions.

Table 7: Results for problem size 35 × 20 × 35.

| | SVRC1 | SVRC2 | SSA1 | SSA2 | SSA3 | SGA1 | SGA2 | SGA3 |
|---|---|---|---|---|---|---|---|---|
| Class I | | | | | | | | |
| Avg_Profit | 58,064,475.00 | **59,724,004.15** | 58,835,952.34 | 58,247,593.08 | 38,238,724.57 | 55,730,804.56 | 57,808,706.54 | 57,087,459.85 |
| Avg%dev | 2.92 | **0.16** | 1.64 | 2.61 | 36.44 | 6.92 | 3.45 | 4.55 |
| Avg_Evals | 1,949,047.20 | **4,509,049.80** | 48,971.00 | 24,684.00 | 19,717.20 | 1,563,423.40 | 52,121.60 | 66,082.00 |
| Avg_Time | 6,733.20 | **3,663.56** | 32.22 | 165.15 | 120.47 | 554.06 | 849.38 | 1,667.85 |
| Class II | | | | | | | | |
| Avg_Profit | 867,540,166.20 | 981,487,545.86 | 981,963,721.23 | 974,060,070.59 | 565,928,537.54 | **984,989,935.75** | 978,690,227.10 | 983,533,295.26 |
| Avg%dev | 12.11 | 0.73 | 0.7 | 1.47 | 42.52 | **0.38** | 1.01 | 0.52 |
| Avg_Evals | 18,860,758.20 | 19,184,769.80 | 829,529.00 | 276,866.20 | 148,021.40 | **35,715,505.40** | 627,077.60 | 637,319.40 |
| Avg_Time | 68,691.96 | 14,677.11 | 337.97 | 5,037.87 | 2,320.08 | **13,272.93** | 16,920.93 | 17,572.59 |
| Class III | | | | | | | | |
| Avg_Profit | 266,985,536.25 | **309,497,478.07** | 303,366,283.17 | 305,752,432.09 | 216,893,443.36 | 308,695,294.99 | 307,980,118.48 | 305,808,477.36 |
| Avg%dev | 13.75 | **0.07** | 2.07 | 1.28 | 29.74 | 0.34 | 0.56 | 1.26 |
| Avg_Evals | 14,534,583.40 | **14,115,188.60** | 464,191.00 | 190,929.40 | 158,399.80 | 19,569,407.20 | 434,557.00 | 212,050.20 |
| Avg_Time | 52,247.29 | **10,590.83** | 332.6 | 3,632.26 | 2,807.67 | 7,096.90 | 8,490.13 | 6,104.69 |

The bold font in the tables refers to the algorithms that reached best solutions.

of business entities with limited capacity. Using the SCND-COQ model can assist firms in improving their profitability and quality simultaneously.

Noteworthy, in the classes where SGA2, SGA3, SSA1, and SSA2 achieve better profits than SVRC2, the average percentage difference between the solutions was less than 0.73%. This may suggest that these methods could be close to the global solution for the classes where the optimal is unknown. As the problem size continues to increase, using SVRC2 will entail considerable computational burden. Thus, the use of GA-based procedures and SSA1 is more attractive for large problems and these methods find solutions close to the best-found solutions. The computational results demonstrate that the state and individual representation in the simulated annealing and the genetic algorithm, respectively, has a significant impact on the solution quality.

A possibility for future research would be to develop additional algorithms specifically designed for the capacitated model; metaheuristics such as SA, the GA, Tabu search, and scatter search could be explored. In addition, a multiproduct and multiobjective capacitated supply chain network design problem including COQ could be studied.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This research was funded in-part by the University of Texas at San Antonio, Office of the Vice President for Research.

## References

[1] L. Li and J. N. Warfield, "Perspectives on quality coordination and assurance in global supply chains," *International Journal of Production Research*, vol. 49, no. 1, pp. 1–4, 2011.

[2] K. Das, "A quality integrated strategic level global supply chain model," *International Journal of Production Research*, vol. 49, no. 1, pp. 5–31, 2011.

[3] S. K. Srivastava, "Towards estimating cost of quality in supply chains," *Total Quality Management and Business Excellence*, vol. 19, no. 3, pp. 193–208, 2008.

[4] A. Ramudhin, C. Alzaman, and A. A. Bulgak, "Incorporating the cost of quality in supply chain design," *Journal of Quality in Maintenance Engineering*, vol. 14, no. 1, pp. 71–86, 2008.

[5] C. Alzaman, A. A. Bulgak, and A. Ramudhin, "Quality in operational supply chain networks: an aerospace case study," *International Journal of Operational Research*, vol. 9, no. 4, pp. 426–442, 2010.

[6] Juran, *Quality Control Handbook*, McGraw-Hill, New York, NY, USA, 1951.

[7] K. K. Castillo-Villar, N. R. Smith, and J. L. Simonton, "A model for supply chain design considering the cost of quality," *Applied Mathematical Modelling*, vol. 36, no. 12, pp. 5920–5935, 2012.

[8] K. K. Castillo-Villar, N. R. Smith, and J. L. Simonton, "The impact of the cost of quality on serial supply chain network design," *International Journal of Production Research*, vol. 50, no. 19, pp. 5544–5566, 2012.

[9] D. P. Ballou and H. L. Pazer, "The impact of inspector fallibility on the inspection policy in serial production systems," *Management Science*, vol. 28, no. 4, pp. 387–399, 1982.

[10] J. T. Godfrey and W. R. Pasewark, "Controlling quality costs," *Management Accounting*, vol. 69, no. 9, pp. 48–51, 1988.

[11] W.-N. Pi and C. Low, "Supplier evaluation and selection using Taguchi loss functions," *International Journal of Advanced Manufacturing Technology*, vol. 26, no. 1-2, pp. 155–160, 2005.

[12] M. R. Bussieck and A. Pruessner, "Mixed-integer nonlinear programming," *SIAG/OPT Newsletter: Views & News*, vol. 14, no. 1, pp. 19–22, 2003.

[13] T. A. Feo and M. G. C. Resende, "Greedy randomized adaptive search procedures," *Journal of Global Optimization*, vol. 6, no. 2, pp. 109–133, 1995.

[14] M. Resende and C. Ribeiro, "Greedy randomized adaptive search procedures," in *Handbook of Metaheuristics*, F. Glover and G. Kochenberger, Eds., pp. 219–249, Springer, New York, NY, USA, 2003.

[15] P. Festa and M. G. C. Resende, "An annotated bibliography of GRASP-part I: algorithms," *International Transactions in Operational Research*, vol. 16, no. 1, pp. 1–24, 2009.

[16] R. W. Eglese, "Simulated annealing: a tool for operational research," *European Journal of Operational Research*, vol. 46, no. 3, pp. 271–281, 1990.

[17] E. Aarts, J. H. M. Korst, and P. J. M. van Laarhoven, "Local search in combinatorial optimization," in *Local Search in Combinatorial Optimization*, John Wiley & Sons, England, 1997.

[18] B. Suman and P. Kumar, "A survey of simulated annealing as a tool for single and multiobjective optimization," *Journal of the Operational Research Society*, vol. 57, no. 10, pp. 1143–1160, 2006.

[19] J. H. Holland, *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, Mich, USA, 1975.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 649471, 6 pages
http://dx.doi.org/10.1155/2014/649471

*Research Article*

# The Solution of Fourth Order Boundary Value Problem Arising out of the Beam-Column Theory Using Adomian Decomposition Method

**Omer Kelesoglu**

*Department of Civil Engineering, Technology Faculty, Firat University, 23119 Elazig, Turkey*

Correspondence should be addressed to Omer Kelesoglu; okelesoglu@firat.edu.tr

Received 23 November 2013; Accepted 2 March 2014; Published 30 March 2014

Academic Editor: Sergii V. Kavun



Adomian decomposition method (ADM) is applied to linear nonhomogeneous boundary value problem arising from the beam-column theory. The obtained results are expressed in tables and graphs. We obtain rapidly converging results to exact solution by using the ADM. This situation indicates that the method is appropriate and reliable for such problems.

## 1. Introduction

It is possible to model many of the physical events that take place in nature using linear and nonlinear differential equations. This modeling enables us to understand and interpret the particular event in a much better manner. Thus, finding the analytical and approximate solutions of such models with initial and boundary conditions gain importance. Differential equations have had an important place in engineering since many years. Scientists and engineers generally examine systems that undergo changes.

Many methods have been developed to determine the analytical and approximate solutions of linear and nonlinear differential equations with initial and boundary value conditions and among these methods, the ADM [1–6], homotopy perturbation method [7–11], variational iteration method [12–19], and homotopy analysis method [20–25] can be listed. Package software such as Matematica, Maple, or Matlab is used to overcome tedious algebraic operations when using these methods. The determination of the analytical and approximate solutions of linear and nonlinear differential equations is an important topic for civil engineering, because these equations are the mathematical models of complex events that occur in engineering.

In this study, the approximate solution of the fourth order linear nonhomogeneous differential equation with initial and boundary conditions that arises in the beam-column theory will be determined via the ADM and comparisons will be made with the existing results in the literature.

The ADM was first put forth in the 1980s by an American scientist named Adomian [26]. This method is based on the decomposition of the unknown function. Using this method it is possible to determine the approximate solutions for the linear and nonlinear ordinary and partial differential equations. The ADM has been used effectively by researchers during 1990 and 2007 especially for the solution of differential and integral equations [27–32]. Using this method a nonlinear problem can be applied directly without discretization and linearization. Hence, this is a method that is preferred by researchers.

In this study, the analytical and approximate solutions of a fourth order linear boundary value problem were calculated using the ADM. In addition, the fourth order linear ordinary differential equation set used in the beam-column theory was solved under specific boundary conditions. The ADM used provides realistic solutions without changing the linear or nonlinear differential equation model. Numerical results close to the real solution can be found by calculating the terms of the limited number of decomposition series. It is possible to find the numerical solution of a differential equation using this method without the necessity of indexing.

## 2. Adomian Decomposition Method

The method has many advantages in comparison with many other traditional methods such as Finite differences, Finite elements, and Galerkin method [33–35]. The method gives converging results adapted to the problems. It is a disadvantage that convergence interval is small and the calculation of the ADM polynomials occurs in nonlinear problems in the ADM.

We now consider second order nonlinear ordinary differential equations with initial conditions. This equation can be written in an operator by

$$Lu + Ru + Nu = f, \tag{1}$$

where $L$ is the lower-order derivative, which is assumed to be invertible, $R$ is a linear differential operator of order greater than $L$, $N$ is a nonlinear differential operator, and $f$ is a source term. We next apply the inverse operator $L^{-1}$ to both sides of (1) and use the given initial conditions to get

$$u(x) = u(0) + xu'(0) + L^{-1}(g) - L^{-1}(Ru) - L^{-1}(Nu), \tag{2}$$

where

$$L = \frac{d^2}{dx^2}, \qquad L^{-1} = \int_0^x \int_0^x dx\,dx. \tag{3}$$

The ADM consists in decomposing the unknown function $u(x)$ of any equation into a sum of infinite number of components given by the decomposition series

$$u(x) = \sum_{n=0}^{\infty} u_n(x), \tag{4}$$

where the components $u_n(x)$, $n \geq 0$ are to be determined in a recursive manner. The nonlinear term $Nu$ will be decomposed by the infinite series of polynomials given by

$$Nu = \sum_{n=0}^{\infty} A_n, \tag{5}$$

where $A_n$ are Adomian polynomials [1, 36]. Substituting (3) and (4) into (2) gives

$$\sum_{n=0}^{\infty} u_n = u_0(x) - L^{-1}R\left(\sum_{n=0}^{\infty} u_n(x) - L^{-1}\sum_{n=0}^{\infty} A_n\right), \tag{6}$$

where

$$u_0(x) = u(0) + xu'(0) + L^{-1}(g). \tag{7}$$

The various components $u_n$ of the solution $u$ can be easily determined by using the recursive relation

$$\begin{aligned} u_0(x) &= u(0) + xu'(0) + L^{-1}g(x), \\ u_{k+1}(x) &= -L^{-1}(Ru_k) - L^{-1}(A_k), \quad k \geq 0. \end{aligned} \tag{8}$$

In order to obtain the numerical solutions of $u(x)$ closed solution function using the decomposition method;

$$\varphi_n = \sum_{k=0}^{n} u_k(x) \quad k \geq 0, \tag{9}$$

being; the term,

$$u(x) = \lim_{n\to\infty} \varphi_n, \tag{10}$$

can be calculated by taking into account the reduction formula (8). In addition, the series solution of the decomposition written as (10) generally yields results that rapidly converge for physical problems. The convergence of the decomposition series has been examined by many researchers in the literature. The convergence of the decomposition series has been examined theoretically by Cherruault [37]. In addition to these studies, Abbaoui and Cherruault have suggested a new approach in determining the convergence of the decomposition series [38]. These authors have determined the convergence of the decomposition series method by giving new conditions.

## 3. The Application of the ADM to the Linear Problem with Boundary Condition That Arises in Beam-Column Theory

*3.1. Problem Definition.* Fourth order differential equations consist of various physical problems that are related to the elastic stability theory. Differential relations should be established between the effects of various cross-section effects in order to understand beam-column problems better.

When a cross section of distance $dx$ shown in Figure 1(b) is taken from a beam-column subject to both the $P$ axial load and the $q$ spread load perpendicular to the axis as shown in Figure 1(a), internal forces arise in the element. When the equilibrium equation in the $y$ direction is written with

$$q = -\frac{dV}{dx}, \tag{11}$$

the following ordinary differential equation is found [39–42]:

$$-V + qdx + (V + dV) = 0. \tag{12}$$

The algebraic sums of the forces acting on both surfaces of the cross-section element are the same due to equilibrium. Consider

$$M + qdx\frac{dx}{2} + (V + dV)\,dx - (M + dM) + P\frac{dy}{dx} = 0. \tag{13}$$

Here, $V$ is the shear force acting on the surface of the element, whereas $M$ is the bending moment that tries to bend the cross-section element.

If rotations are assumed to be small and the second order terms in terms of $dx$ are neglected, then (13) becomes

$$V = \frac{dM}{dx} - P\frac{dy}{dx}. \tag{14}$$

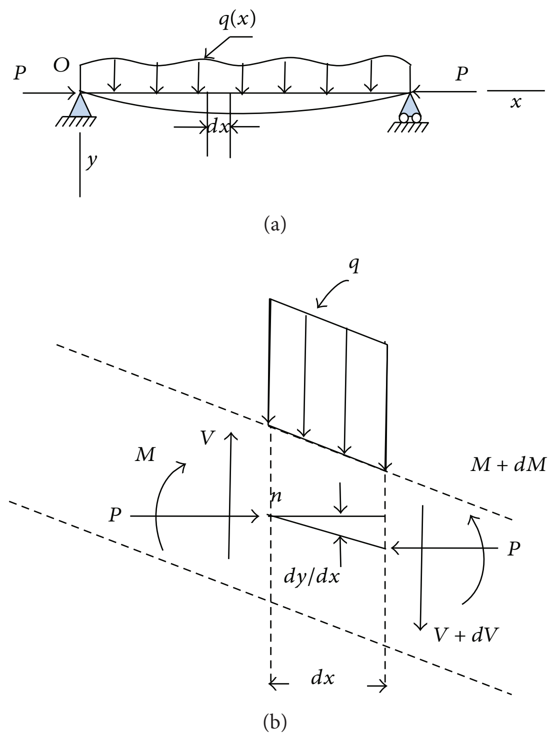



Figure 1: Cross-sectional analysis of thecolumn-beam element.

Since rotations are assumed to be small, if $d^2/dx^2 = -M/EI$, then (14) becomes

$$-V = EI\frac{d^3 y}{dx^3} + P\frac{dy}{dx}. \tag{15}$$

Here, $EI$ represents bending rigidity. If the derivative of both sides of (15) is taken in terms of $x$, then the fourth order linear differential equation for the elastic curve is found as such:

$$EI\frac{d^4 y}{dx^4} + P\frac{d^2 y}{dx^2} = q(x). \tag{16}$$

*3.2. Application of the ADM to the Problem*

*Example 1.* We consider the fourth order linear nonhomogeneous differential equation [43]:

$$\frac{d^4 y}{dx^4} - 2\frac{d^2 y}{dx^2} + y = -8e^x, \quad x \in [0,1], \tag{17}$$

with the boundary conditions of

$$y(0) = y''(0) = 1, \qquad y'(1) = y''(1) = -e. \tag{18}$$

If (18) is written out in operator form, we obtain

$$Ly = 2y'' - y - 8e^x. \tag{19}$$

Here,

$$L = \frac{d^4}{dx^4}, \qquad L^{-1}(\cdot) = \int_0^x \int_0^x \int_0^x \int_0^x (\cdot)\, dx\, dx\, dx\, dx \tag{20}$$

are the derivative and integral operators. If the $L^{-1}$ inverse operator is applied to (19) and initial conditions are taken, we find

$$y(x) = Ax + \frac{1}{3!}Bx^3 - 8L^{-1}(e^x) + 2L^{-1}(y'') - L^{-1}(y), \tag{21}$$

where $A = y'(0)$ and $B = y'''(0)$. If we use (3) in (21), then we find

$$\begin{aligned}\sum_{n=0}^{\infty} y_n(x) &= Ax + \frac{1}{3!}Bx^3 - 8L^{-1}e^x \\ &\quad + 2L^{-1}\left(\sum_{n=0}^{\infty} y_n''(x)\right) - L^{-1}\left(\sum_{n=0}^{\infty} y_n(x)\right),\end{aligned} \tag{22}$$

whereas the reduction formula given below can be written using (22):

$$\begin{aligned}y_0(x) &= Ax + \frac{1}{3!}Bx^3 - 8L^{-1}(e^x) \\ y_{n+1}(x) &= 2L^{-1}(y_n'') - L^{-1}(y_n) \\ &\qquad n \ge 0.\end{aligned} \tag{23}$$

The $A$ and $B$ constants in the reduction formula will be determined using boundary conditions (18) after finding the decomposition series. From the reduction relation, we can obtain the solution terms of the decomposition series as

$$\begin{aligned}y_0(x) &= 8 - 8e^x + (8 + A)x + 4x^2 + \frac{1}{6}(8 + B)x^3, \\ y_1(x) &= 8 - 8e^x + 8x + 4x^2 + \frac{4x^3}{3} + \frac{x^4}{3} \\ &\quad - \frac{1}{120}(-8 + A - 2B)x^5 - \frac{x^6}{90} - \frac{(8 + B)x^7}{5040} \\ y_2(x) &= 8 - 8e^x + 8x + 4x^2 + \frac{4x^3}{3} + \frac{x^4}{3} - \frac{x^5}{15} + \frac{x^6}{90} \\ &\quad - \frac{(A - 2(2 + B))x^7}{2520} - \frac{x^8}{1680} \\ &\quad + \frac{(A - 4(6 + B))x^9}{362880} + \frac{x^{10}}{453600} + \frac{(8 + B)x^{11}}{39916800}, \\ y_3(x) &= 8 - 8e^x + 8x + 4x^2 + \frac{4x^3}{3} + \frac{x^4}{3} + \frac{x^5}{15} \\ &\quad + \frac{x^6}{90} + \frac{x^7}{630} + \frac{x^8}{5040} - \frac{(-2 + A - 2B)x^9}{90720} \\ &\quad - \frac{x^{10}}{64800} + \frac{(-14 + A - 3B)x^{11}}{9979200}\end{aligned}$$

$$+\frac{x^{12}}{11975040}-\frac{(-40+A-6B)\,x^{13}}{6227020800}-\frac{x^{14}}{10897286400}-\frac{(8+B)\,x^{15}}{1307674368000},$$

$$\begin{aligned}y_4(x) &= 8-8e^{x}+8x+4x^2+\frac{4x^3}{3}+\frac{x^4}{3}+\frac{x^5}{15}\\&+\frac{x^6}{90}+\frac{x^7}{630}+\frac{x^8}{5040}+\frac{x^9}{45360}\\&+\frac{x^{10}}{453600}-\frac{(-1+A-2B)\,x^{11}}{4989600}-\frac{x^{12}}{3991680}\\&+\frac{(-30+3A-8B)\,x^{13}}{1556755200}+\frac{17x^{14}}{10897286400}\\&-\frac{(-6B+3A-12B)\,x^{15}}{653837184000}-\frac{x^{16}}{373621248000}\\&+\frac{(A-8(7+B))\,x^{17}}{355687428096000}+\frac{x^{18}}{800296713216000}\\&+\frac{(8+B)\,x^{19}}{121645100408832000}.\end{aligned} \tag{24}$$

If these terms are placed in (4), we obtain the approximate solution obtained via the ADM using the five terms of (17) and problem (18) can be written as

$$y(x)=\sum_{i=0}^{4}y_i(x)=y_0(x)+y_1(x)+y_2(x)+y_3(x)+y_4(x), \tag{25}$$

$$\begin{aligned}y(x) &= 40-40e^{x}+(40+A)\,x+20x^2+\frac{1}{6}(40+B)\,x^3\\&+\frac{4x^4}{3}-\frac{1}{120}(A-2(16+B))\,x^5+\frac{x^6}{45}\\&-\frac{(-16+2A-3B)\,x^7}{5040}-\frac{x^8}{5040}\\&-\frac{(8+3A-4B)\,x^9}{362880}-\frac{x^{10}}{90720}-\frac{(40+4A-5B)\,x^{11}}{39916800}\\&-\frac{x^{12}}{5987520}+\frac{(-80+11A-26B)\,x^{13}}{6227020800}+\frac{x^{14}}{681080400}\\&-\frac{(-128+6A-23B)\,x^{15}}{1307674368000}-\frac{x^{16}}{373621248000}\\&+\frac{(A-8(7+B))\,x^{17}}{355687428096000}+\frac{x^{18}}{800296713216000}\\&+\frac{(8+B)\,x^{19}}{121645100408832000}.\end{aligned} \tag{26}$$

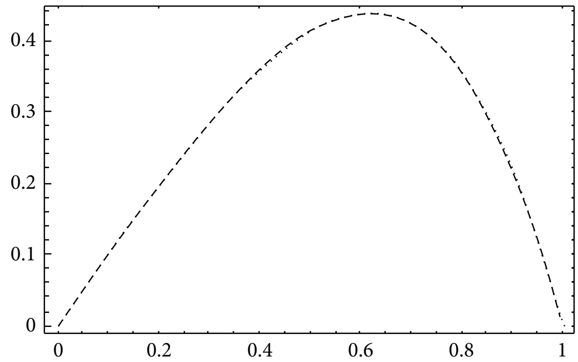



FIGURE 2: Graph showing the ADM and analytical solution obtained using the five terms of the decomposition series.

TABLE 1: Numerical results obtained using the five terms of the decomposition series.

| $x$ | Analytical solution $y(x)$ | Approximate solution $\emptyset_5$ | Error $\lvert y(x)-\emptyset_5\rvert$ |
|---|---|---|---|
| 0 | 0. | 0. | 0. |
| 0.1 | 0.099465 | 0.099465 | $2.92821\times10^{-15}$ |
| 0.2 | 0.195424 | 0.195424 | $-2.16493\times10^{-15}$ |
| 0.3 | 0.283470 | 0.283470 | $6.10623\times10^{-16}$ |
| 0.4 | 0.358038 | 0.358038 | $-1.30451\times10^{-13}$ |
| 0.5 | 0.412180 | 0.412180 | $-2.37055\times10^{-12}$ |
| 0.6 | 0.437309 | 0.437309 | $-2.58711\times10^{-11}$ |
| 0.7 | 0.422888 | 0.422888 | $-1.94957\times10^{-10}$ |
| 0.8 | 0.356087 | 0.356087 | $-1.12373\times10^{-9}$ |
| 0.9 | 0.221364 | 0.221364 | $-5.27758\times10^{-9}$ |
| 1.0 | **0.** | $\mathbf{2.10884\times10^{-8}}$ | $\mathbf{-2.10884\times10^{-8}}$ |

If we use boundary conditions (18) in solutions series (26), then we have $A=1$ and $B=-3$. The analytical solution of problem (17) and (18) is [43]

$$y(x)=x(1-x)\,e^{x}. \tag{27}$$

The numerical results and graphs obtained for five terms using the solution series in (26) have been given below.

## 4. Results and Discussion

In this study, the approximate solution of the fourth order boundary value problem arising in beam-column theory has been determined using the ADM. This method can be applied on differential equations without the need for discretization, indexing, or linearization.

Nonhomogeneous problem has been handled regarding the topic and the obtained results have been given in Table 1 and Figure 2. As can be seen from the table and figure, the method used has given results that converge rapidly to the analytical solution with the removal of a few terms from the

TABLE 2: Comparison of the absolute error obtained using the spline method (SM) and the ADM.

| $x$ | SM $h = 1/5$ | SM $h = 1/10$ | ADM $\emptyset_5$ | CPU time |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0.542 |
| 0.2 | $1.960E-5$ | $1.228E-6$ | $-2.164E-15$ | 0.846 |
| 0.4 | $3.211E-5$ | $2.012E-6$ | $-1.304E-13$ | 0.870 |
| 0.6 | $3.568E-5$ | $2.235E-6$ | $-2.587E-11$ | 0.875 |
| 0.8 | $2.683E-5$ | $1.681E-6$ | $-1.123E-9$ | 0.891 |
| 1 | 0 | 0 | $-2.108E-8$ | 0.902 |

solution series. This shows that the method is suitable and reliable for such problems. In addition, the handled example has been compared with the results obtained by Chen and Atsuta [41, 42] using nonpolynomial spline method. As seen in Table 2 values close to those obtained in this study were found only in case when the "$h$" step was small. New terms can be added to the solution series of the ADM to obtain results that are much better than those of the spline method.

In Figure 2, five terms of the decomposition method have been taken from the analytical solution and decomposition series of example have been drawn in two dimensional graphs. As seen in these graphs, the analytical and approximate solutions cannot be distinguished.

In conclusion, it has been determined that the ADM can be applied to the linear homogeneous and nonhomogeneous boundary value problems that arise in civil engineering in the beam-column theory. Solutions that converge rapidly to the analytical solution can be found without changing the nature of the physical phenomenon. In addition, the calculations for this method can be carried out using software such as Mathematica, Maple, and Matlab.

## Conflict of Interests

The author declares that there is no conflict of interests.

## References


[1] G. Adomian, *Solving Frontier Problems of Physics: The Decomposition Method*, Kluwer Academic Publishers, Boston, Mass, USA, 1994.

[2] G. Adomian, "A review of the decomposition method and some recent results for nonlinear equations," *Mathematical and Computer Modelling*, vol. 13, no. 7, pp. 17–43, 1990.

[3] G. Adomian and R. Rach, "Equality of partial solutions in the decomposition method for linear or nonlinear partial differential equations," *Computers & Mathematics with Applications*, vol. 19, no. 12, pp. 9–12, 1990.

[4] M. Inc, "On numerical solutions of partial differential equations by the decomposition method," *Kragujevac Journal of Mathematics*, vol. 26, pp. 153–164, 2004.

[5] M. Inc, "Decomposition method for solving parabolic equations in finite domains," *Journal of Zhejiang University SCIENCE A*, vol. 6, no. 10, pp. 1058–1064, 2005.

[6] M. Inc, Y. Cherruault, and K. Abbaoui, "A computational approach to the wave equations: an application of the decomposition method," *Kybernetes*, vol. 33, no. 1, pp. 80–97, 2004.

[7] Z. M. Odibat, "A new modification of the homotopy perturbation method for linear and nonlinear operators," *Applied Mathematics and Computation*, vol. 189, no. 1, pp. 746–753, 2007.

[8] S. R. S. Alizadeh, G. G. Domairry, and S. Karimpour, "An approximation of the analytical solution of the linear and nonlinear integro-differential equations by homotopy perturbation method," *Acta Applicandae Mathematicae*, vol. 104, no. 3, pp. 355–366, 2008.

[9] Y.-G. Wang, H.-F. Song, and D. Li, "Solving two-point boundary value problems using combined homotopy perturbation method and Green's function method," *Applied Mathematics and Computation*, vol. 212, no. 2, pp. 366–376, 2009.

[10] J. Biazar and H. Ghazvini, "Convergence of the homotopy perturbation method for partial differential equations," *Nonlinear Analysis: Real World Applications*, vol. 10, no. 5, pp. 2633–2640, 2009.

[11] C.-S. Liu, "The essence of the homotopy analysis method," *Applied Mathematics and Computation*, vol. 216, no. 4, pp. 1299–1303, 2010.

[12] S. Momani and S. Abuasad, "Application of He's variational iteration method to Helmholtz equation," *Chaos, Solitons & Fractals*, vol. 27, no. 5, pp. 1119–1123, 2006.

[13] E. M. Abulwafa, M. A. Abdou, and A. A. Mahmoud, "Nonlinear fluid flows in pipe-like domain problem using variational-iteration method," *Chaos, Solitons & Fractals*, vol. 32, no. 4, pp. 1384–1397, 2007.

[14] N. H. Sweilam and M. M. Khader, "Variational iteration method for one dimensional nonlinear thermoelasticity," *Chaos, Solitons & Fractals*, vol. 32, no. 1, pp. 145–149, 2007.

[15] L. Xu, "Variational iteration method for solving integral equations," *Computers & Mathematics with Applications*, vol. 54, no. 7-8, pp. 1071–1078, 2007.

[16] J.-H. He, A.-M. Wazwaz, and L. Xu, "The variational iteration method: reliable, efficient, and promising," *Computers & Mathematics with Applications*, vol. 54, no. 7-8, pp. 879–880, 2007.

[17] L. Xu, J.-H. He, and A.-M. Wazwaz, "Preface variational iteration method—reality, potential, and challenges," *Journal of Computational and Applied Mathematics*, vol. 207, no. 1, pp. 1–2, 2007.

[18] S. B. Coşkun and M. T. Atay, "Analysis of convective straight and radial fins with temperature-dependent thermal conductivity using variational iteration method with comparison with respect to finite element analysis," *Mathematical Problems in Engineering*, vol. 2007, Article ID 42072, 15 pages, 2007.

[19] M. T. Atay and S. B. Coşkun, "Effects of nonlinearity on the variational iteration solutions of nonlinear two-point boundary value problems with comparison with respect to finite element analysis," *Mathematical Problems in Engineering*, vol. 2008, Article ID 857296, 10 pages, 2008.

[20] M. Inc and Y. Uğurlu, "Numerical simulation of the regularized long wave equation by He's homotopy perturbation method," *Physics Letters A: General, Atomic and Solid State Physics*, vol. 369, no. 3, pp. 173–179, 2007.

[21] A. S. Bataineh, M. S. M. Noorani, and I. Hashim, "Homotopy analysis method for singular IVPs of Emden-Fowler type," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 4, pp. 1121–1131, 2009.

[22] S. Abbasbandy, E. Babolian, and M. Ashtiani, "Numerical solution of the generalized Zakharov equation by homotopy analysis method," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 12, pp. 4114–4121, 2009.

[23] M. M. Rashidi and S. Dinarvand, "Purely analytic approximate solutions for steady three-dimensional problem of condensation film on inclined rotating disk by homotopy analysis method," *Nonlinear Analysis: Real World Applications*, vol. 10, no. 4, pp. 2346–2356, 2009.

[24] L. Song and H. Zhang, "Solving the fractional BBM-Burgers equation using the homotopy analysis method," *Chaos, Solitons & Fractals*, vol. 40, no. 4, pp. 1616–1622, 2009.

[25] S. Liao, "On the relationship between the homotopy analysis method and Euler transform," *Communications in Nonlinear Science and Numerical Simulation*, vol. 15, no. 6, pp. 1421–1431, 2010.

[26] G. Adomian, *Nonlinear Stochastic Operator Equations*, Academic Press, San Diego, Calif, USA, 1986.

[27] M. Inc and M. Işık, "Adomian decomposition method for three-dimensional parabolic equation with non-classic boundary conditions," *Journal of Analysis*, vol. 11, pp. 43–51, 2003.

[28] F. Abdelwahid, "A mathematical model of Adomian polynomials," *Applied Mathematics and Computation*, vol. 141, no. 2-3, pp. 447–453, 2003.

[29] E. Babolian and S. Javadi, "New method for calculating Adomian polynomials," *Applied Mathematics and Computation*, vol. 153, no. 1, pp. 253–259, 2004.

[30] Y. Cherruault, M. Inc, and K. Abbaoui, "On the solution of the non-linear Korteweg-de Vries equation by the decomposition method," *Kybernetes*, vol. 31, no. 5, pp. 766–772, 2002.

[31] A. M. A. El-Sayed and M. Gaber, "The Adomian decomposition method for solving partial differential equations of fractal order in finite domains," *Physics Letters A*, vol. 359, no. 3, pp. 175–182, 2006.

[32] E. Momoniat, T. A. Selway, and K. Jina, "Analysis of Adomian decomposition applied to a third-order ordinary differential equation from thin film flow," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 66, no. 10, pp. 2315–2324, 2007.

[33] L. L. Thompson and P. M. Pinsky, "A Galerkin least-squares finite element method for the two-dimensional Helmholtz equation," *International Journal for Numerical Methods in Engineering*, vol. 38, no. 3, pp. 371–397, 1995.

[34] J. Dolbow and T. Belytschko, "Numerical integration of the Galerkin weak form in meshfree methods," *Computational Mechanics*, vol. 23, no. 3, pp. 219–230, 1999.

[35] S. N. Atluri and T. Zhu, "A new meshless local Petrov-Galerkin (MLPG) approach in computational mechanics," *Computational Mechanics*, vol. 22, no. 2, pp. 117–127, 1998.

[36] A. M. Wazwaz, *Partial Differential Equations: Methods and Applications*, Balkema Publishers, Rotterdam, The Netherlands, 2002.

[37] Y. Cherruault, "Convergence of Adomian's method," *Kybernetes*, vol. 18, no. 2, pp. 31–38, 1989.

[38] K. Abbaoui and Y. Cherruault, "New ideas for proving convergence of decomposition methods," *Computers & Mathematics with Applications*, vol. 29, no. 7, pp. 103–108, 1995.

[39] S. P. Timoshenko and J.M. Gere, *Theory of Elastic Stability*, McGraw-Hill, New York, NY, USA, 2nd edition, 1961.

[40] S. P. Timoshenko and J.M. Gere, *Theory of Elastic Stability*, McGraw-Hill, New York, NY, USA, 1985.

[41] W. F. Chen and T. Atsuta, *Theory of Beam-Columns*, vol. 1 of *Inplane Behavior and Design*, McGraw-Hill, New York, NY, USA, 1976.

[42] W. F. Chen and T. Atsuta, *Theory ofBeam-Columns*, vol. 2 of *Space Behavior and Design*, McGraw-Hill, New York, NY, USA, 1977.

[43] O. A. Taiwo and O. M. Ogunlaran, "A non-polynomial spline method for solving linear fourth-order boundary-value problems," *International Journal of Physical Sciences*, vol. 6, no. 13, pp. 3246–3254, 2011.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 317697, 6 pages
http://dx.doi.org/10.1155/2014/317697

*Research Article*

# Integer Programming Formulations for Approximate Packing Circles in a Rectangular Container

**Igor Litvinchev[1,2] and Edith Lucero Ozuna Espinosa[2]**

[1] *Complex Systems Department, Computing Center, Russian Academy of Sciences, 119333 Moscow, Russia*
[2] *Faculty of Mechanical and Electrical Engineering, Nuevo Leon State University, 66450 Monterrey, NL, Mexico*

Correspondence should be addressed to Igor Litvinchev; litvin@ccas.ru

Received 23 December 2013; Accepted 15 February 2014; Published 25 March 2014

Academic Editor: Vyacheslav Kalashnikov



A problem of packing a limited number of unequal circles in a fixed size rectangular container is considered. The aim is to maximize the (weighted) number of circles placed into the container or minimize the waste. This problem has numerous applications in logistics, including production and packing for the textile, apparel, naval, automobile, aerospace, and food industries. Frequently the problem is formulated as a nonconvex continuous optimization problem which is solved by heuristic techniques combined with local search procedures. New formulations are proposed for approximate solution of packing problem. The container is approximated by a regular grid and the nodes of the grid are considered as potential positions for assigning centers of the circles. The packing problem is then stated as a large scale linear 0-1 optimization problem. The binary variables represent the assignment of centers to the nodes of the grid. Nesting circles inside one another is also considered. The resulting binary problem is then solved by commercial software. Numerical results are presented to demonstrate the efficiency of the proposed approach and compared with known results.

## 1. Introduction

Packing problems generally consist of packing of a certain number of items of known dimensions into one or more large objects or containers so as to minimize a certain objective (e.g., the unused part of the objects or waste). The shape of items and containers may vary from a circle, a square, a rectangle, and so forth.

This problem has been applied in different areas, such as the coverage of a geographical area with cell transmitters, storage of a cylindrical drums into containers or stocking them into an open area, packaging bottles or cans into the smallest box, planting trees in a given region so as to maximize the forest density and the distance between the trees, and so forth [1–3]. One can find other applications in the motor cycle industry, circular cutting, communication networks, facility location, and dashboard layout [4–8].

In this paper we address the problem of packing a set of circular items in a rectangular container. There are two principal types of objectives that have been used in the literature: (a) regard the circles (not necessary equal) as being of fixed size and the container as being of variable size and (b) regard the circles and the container as being of fixed size and minimize "waste."

Examples of the first approach include the following [9].

(i) For the square container minimize the length of the side and hence minimize the perimeter and area of the square.

(ii) Minimize the perimeter of the rectangle.

(iii) Minimize the area of the rectangle.

(iv) Considering one dimension of the rectangle as fixed, minimize the other dimension. Problems of this type are often referred to as strip packing problems (or as circular open dimension problems).

For the second approach various definitions of the waste can be used. The waste can be defined in relation to circles not packed (e.g., the number of unpacked circles or the perimeter/area of unpacked circles) or introducing a value associated with each circle that is packed (e.g., area of the circles packed), and so forth.

Many variants of packing circular objects in the plane have been formulated as nonconvex (continuous) optimization problems with decision variables being coordinates of the centres. The nonconvexity is mainly provided by no overlapping conditions between circles. These conditions typically state that the Euclidean distance separating the centres of the circles is greater than a sum of their radii. The nonconvex problems can be tackled by available nonlinear programming (NLP) solvers; however, most NLP solvers fail to identify global optima. Thus, the nonconvex formulation of circular packing problem requires algorithms which mix local searches with heuristic procedures in order to widely explore the search space. It is impossible to give a detailed overview on the existing solution strategies and numerical results within the framework of a single short paper. We will refer the reader to review papers presenting the scope of techniques and applications for the circle packing problem (see, e.g., [1, 2, 10–13] and the references therein).

In this paper we propose a new formulation for approximate solution of packing problems based on using a regular grid to approximate the container. The nodes of the grid are considered as potential positions for assigning centers of the circles. The packing problem is then stated as a large scale linear 0-1 optimization problem. The binary variables represent the assignment of centers to the nodes of the grid. The resulting binary problem is then solved by commercial software. To the best of our knowledge, the idea to use a grid was first implemented by Beasley [14] in the context of cutting problems. Recently this approach was applied in [15, 16] for packing problems. This work is a continuation of [16].

## 2. The Model

Suppose we have nonidentical circles $C_k$ of known radius $R_k$, $k \in K = \{1, 2, \ldots K\}$. Let at most $M_k$ circles $C_k$ be available for packing, and at least $m_k$ of them have to be packed. Denote by $i \in I = \{1, 2 \ldots, n\}$ the node points of a regular grid covering the rectangular container. Let $F \subseteq I$ be the grid points lying on the boundary of the container. Denote by $d_{ij}$ the Euclidean distance between points $i$ and $j$ of the grid. Define binary variables $x_i^k = 1$ if centre of a circle $C_k$ is assigned to the point $i$; $x_i^k = 0$, otherwise.

In what follows we will distinguish two cases of circle packing, depending on whether nesting circles inside one another is permitted or not. To the best of our knowledge, nesting problem was first mentioned in [6] in the context of packing pipes of different diameters into a shipping container. Compared to the standard packing, packing with nesting is much less investigated.

Consider first the problem without nesting. In order to let the circle $C_k$ assigned to the point $i$ be nonoverlapping with other circles being packed, it is necessary that $x_j^l = 0$ for $j \in I$, $l \in K$, such that $d_{ij} < R_k + R_l$. For fixed $i$, $k$ let $N_{ik} = \{j, l : i \neq j,\ d_{ij} < R_k + R_l\}$. Let $n_{ik}$ be the cardinality of $N_{ik} : n_{ik} = |N_{ik}|$. Then the problem of maximizing the area covered by the circles can be stated as follows:

$$\max \sum_{i \in I} \sum_{k \in K} R_k^2 x_i^k, \tag{1}$$

$$\text{subject to} \quad m_k \leq \sum_{i \in I} x_i^k \leq M_k, \quad k \in K, \tag{2}$$

$$\sum_{k \in K} x_i^k \leq 1, \quad i \in I \setminus F, \tag{3}$$

$$R_k x_i^k \leq \min_{j \in F} d_{ij}, \quad i \in I,\ k \in K, \tag{4}$$

$$\begin{gathered} x_i^k + x_j^l \leq 1, \\ \text{for } i \in I,\ k \in K,\ (j, l) \in N_{ik} \end{gathered} \tag{5}$$

$$x_i^k \in \{0, 1\}, \quad i \in I,\ k \in K. \tag{6}$$

Constraints (2) ensure that the number of circles packed is between $m_k$ and $M_k$; constraints (3) ensure that at most one centre is assigned to any grid point; constraints (4) ensure that the point $i$ cannot be a centre of the circle $C_k$ if the distance from $i$ to the boundary is less than $R_k$; pairwise constraints (5) guarantee that there is no overlapping between the circles; constraints (6) represent the binary nature of variables.

Similar to plant location problems [17] we can state nonoverlapping conditions in a more compact form. Summing up pairwise constraints (5) over $(j, l) \in N_{ik}$ we get

$$n_{ik} x_i^k + \sum_{j,l \in N_{ik}} x_j^l \leq n_{ik} \quad \text{for } i \in I,\ k \in K. \tag{7}$$

Note that constraints similar to (7) were used in [15] for packing equal circles.

**Proposition 1.** *Constraints* (5) *and* (6) *are equivalent to constraints* (6) *and* (7).

*Proof.* If (5) are fulfilled, then obviously (7) hold by construction. If (7) are fulfilled, then for $x_i^k = 1$ we have $\sum_{j,l \in N_{ik}} x_j^l \leq 0$ and hence $x_j^l = 0$ for $j, l \in N_{ik}$ as in (5). If $x_i^k = 0$ in (7), then $\sum_{j,l \in N_{ik}} x_j^l \leq n_{ik}$ holds for all $x_j^l \in \{0, 1\}$. □

Thus the problem (1)–(6) is equivalent to the problem (1)–(4), (6), and (7). To compare two equivalent formulations, let

$$P_1 = \left\{x \geq 0 : x_i^k + x_j^l \leq 1,\ \text{for } i \in I,\ k \in K,\ (j, l) \in N_{ik}\right\},$$

$$P_2 = \left\{x \geq 0 : n_{ik} x_i^k + \sum_{j,l \in N_{ik}} x_j^l \leq n_{ik} \text{ for } i \in I,\ k \in K\right\}. \tag{8}$$

**Proposition 2.** $P_1 \subset P_2$.

*Proof.* Since constraints of $P_2$ are obtained by summing up some constraints of $P_1$, then $P_1 \subseteq P_2$. To show that $P_1 \subset P_2$ we need to find a point in $P_2$ that is not in $P_1$.

Table 1: Results of the numerical experiments.

| Circle radius | $\Delta$ | Problem dimension | Circle number | Complete | Half | Compact | Compact half |
|---|---|---|---|---|---|---|---|
| 0.625 | 0.15625 | 1403 | 10 | 6.4 | 125.5 | 144.8 | 110.6 |
| 0.5625 | 0.0703125 | 2449 | 13 | 50.4 | 647.0 | 5034.5 | 1890.6 |
| 0.5 | 0.125 | 697 | 18 | 2.6 | 3.2 | 5.2 | 24.1 |
| 0.4375 | 0.0546875 | 3666 | 21 | 849.2 | 310.8 | 7459.9 | 3690.5 |
| 0.375 | 0.046875 | 1425 | 32 | 50.0 | 21.4 | 3873.5 | 847.7 |
| 0.3125 | 0.078125 | 2139 | 45 | 403.8 | 183.9 | 6514.3 | 4451.8 |
| 0.275 | 0.06875 | 2880 | 61 | 1032.4 | 415.3 | * | 6985.4 |
| 0.25 | 0.0625 | 3649 | 74 | 1234.8 | 535.5 | 14645.9 | 7840.6 |
| 0.1875 | 0.046875 | 6897 | 140 | 1427.3 | 725.9 | * | 8765.2 |

This point can be constructed as follows. By the definition, $N_{ik} = \{j,l : i \neq j,\ d_{ij} < R_k + R_l\}$ and hence if $(j,l) \in N_{ik}$, then $(i,k) \in N_{jl}$. Choose $(i,k)$ and $(j,l) \in N_{ik}$ such that $n_{ik}, n_{jl} \geq 2$. Set to zero all the variables except $x_i^k$, $x_j^l$. Obviously all constraints (5) corresponding to zero variables are fulfilled. Define $x_i^k, x_j^l$ to fulfil the two remaining constraints as equalities:

$$n_i^k x_i^k + x_j^l = n_i^k, \qquad n_j^l x_j^l + x_i^k = n_j^l \tag{9}$$

with $n_{ik}, n_{jl} \geq 2$. The corresponding solution is

$$x_i^k = \frac{n_{jl}(n_{ik}-1)}{n_{il}n_{ik}-1} < 1, \qquad x_j^l = \frac{n_{ik}(n_{jl}-1)}{n_{il}n_{ik}-1} < 1 \tag{10}$$

with

$$x_i^k + x_j^l = 1 + \frac{1 + n_{jl}n_{ik} - n_{jl} - n_{ik}}{n_{il}n_{ik}-1} > 1. \tag{11}$$

Thus this point violates corresponding constraint (5) in $P_1$ and hence $P_1 \subset P_2$ as desired. □

As follows from Proposition 2, the pairwise formulation (1)–(6) is stronger than the compact one [17].

By the definition, $N_{ik} = \{j,l : i \neq j,\ d_{ij} < R_k + R_l\}$ and hence if $(j,l) \in N_{ik}$, then $(i,k) \in N_{jl}$. Thus a half of the constraints in (5) are redundant:

$$\begin{aligned} x_i^k + x_j^l &\leq 1, \quad \text{for } i \in I,\ k \in K,\ (j,l) \in N_{ik}, \\ x_j^l + x_i^k &\leq 1, \quad \text{for } j \in I,\ l \in K,\ (i,k) \in N_{jl}. \end{aligned} \tag{12}$$

The redundant constraints can be eliminated giving a reduced pairwise nonoverlapping formulation. The overall set of the reduced constraints is independent of whether we will eliminate the first constraint above or the second. However, the way of eliminating redundant constraints will affect the corresponding compact formulation obtained by summing up the reduced constraints.

To consider nesting circles inside one another, we only need to modify the nonoverlapping constraints. In order to let the circle $C_k$ assigned to the point $i$ be nonoverlapping with other circles being packed (including circles places inside this circle), it is necessary that $x_j^l = 0$ for $j \in I, l \in K$, such that $R_k - R_l < d_{ij} < R_k + R_l$. Note that the later condition is always fulfilled for $R_k < R_l$ ($d_{ij} \geq 0$), such that only smaller circles can be placed inside a given circle. For fixed $i$, $k$ let $\Omega_{ik} = \{j,l : i \neq j,\ |R_k - R_l| < d_{ij} < R_k + R_l\}$. Then the problem of packing circles with nesting can be stated as follows:

$$\begin{aligned} &\max \sum_{i \in I} \sum_{k \in K} w_i^k x_i^k \\ \text{subject to} \quad & m_k \leq \sum_{i \in I} x_i^k \leq M_k, \quad k \in K, \\ & R_k x_i^k \leq \min_{j \in F} d_{ij}, \quad i \in I,\ k \in K, \\ & x_i^k + x_j^l \leq 1, \quad \text{for } i \in I,\ k \in K,\ (j,l) \in \Omega_{ik}. \end{aligned} \tag{13}$$

Here weighting coefficients $w_i^k$ may be associated with the area of circles and/or represent the relative importance of subsets of the container.

## 3. Computational Results

A rectangular uniform grid was used in the numerical experiments, such that all grid points are defined by the grid points on its edges. Let $L$ be a horizontal dimension (length) and let $W$ be a vertical dimension (width) of the container; let $M$ be a number of the equidistant grid points on the horizontal edge of the container, while $N$ is a number of the equidistant grid points on its vertical edge. Hence the grid has $M \times N$ node points ($n = M \times N$).

All optimization problems were solved by the system CPLEX 12.5. The runs were executed on a DELL Power Edge T410, Intel Xeon 2.53 Ghz and 16 Gb RAM.

First, we compare different formulations for the case of packing equal circles. The same set of 9 instances as in [15, Table 3] was used for the experiment. The number of circles available for packing was not limited. The following formulations were considered: pairwise formulation (1)–(6) (*complete*), reduced formulation (1)–(6) without redundant constraints (*half*), compact formulation (1)–(4), (6), and (7) (*compact*), and compact formulation obtained from (1)–(6) without redundant constraints (*compact half*). The results of the numerical experiment are presented in Table 1. Here the first four columns present circle radius, dimension of the quadratic cell used to construct the grid ($\Delta$), the number

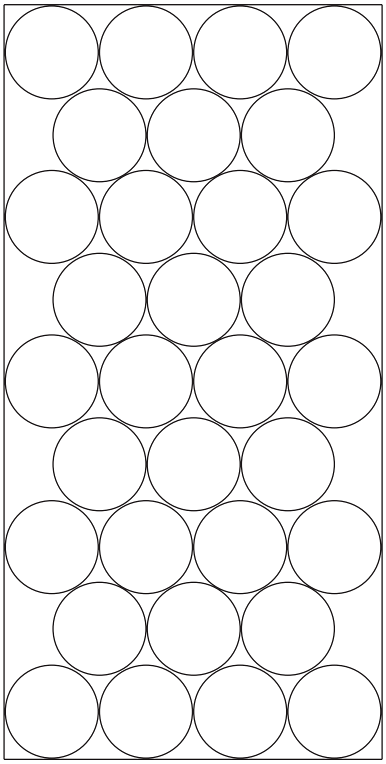


Figure 1: $R = 0.375$.

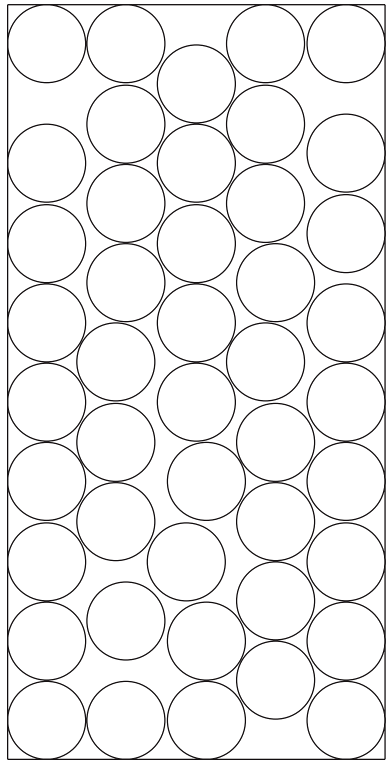


Figure 2: $R = 0.3125$.

of grid points ($n$), and the number of circles packed. For all formulations the same number of circles packed was achieved. The last four columns give CPU time for different formulations (in seconds). For all problem instances *mipgap* = 0 was set for running CPLEX. Asterisks in the column indicate that computation was interrupted after the computational time exceeded 15-hour CPU time.

As we can see from Table 1, CPU time for the complete formulation is typically much lower than for the compact formulation, especially for large problems. Note that according to Proposition 2, complete (pairwise) formulation is stronger than the compact one. Eliminating redundant constraints typically (but not always) reduces CPU time, although for complete formulation eliminating redundancy does not change the continuous relaxation. This effect is well known for users of CPLEX since simply rearranging constraints may result in increase/decrease of computational time, depending on the path selected by branch and bound algorithm. However, we may conclude that eliminating redundant constraints is useful for large problems.

Figures 1, 2, 3, and 4 present packing pictures, for instance, from Table 1 with radii 0.375, 0.3125, 0.4375, and 0.275, respectively.

Figures 5 and 6 present the packing pictures obtained for the container with $L = W = 30$ and $R_1 = 0.6$, $R_2 = 3$, and $R_3 = 6$ using pairwise formulation (1)–(6). We used $M = N = 31$ for Figure 5 and $M = N = 51$ for Figure 6. The solution presented in Figure 5 was obtained in 53.239 sec with *mipgap* = 0, while for Figure 6 the computation was

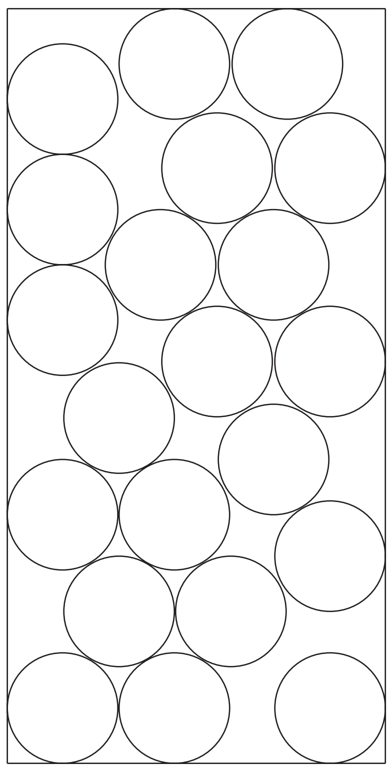


Figure 3: $R = 0.4375$.

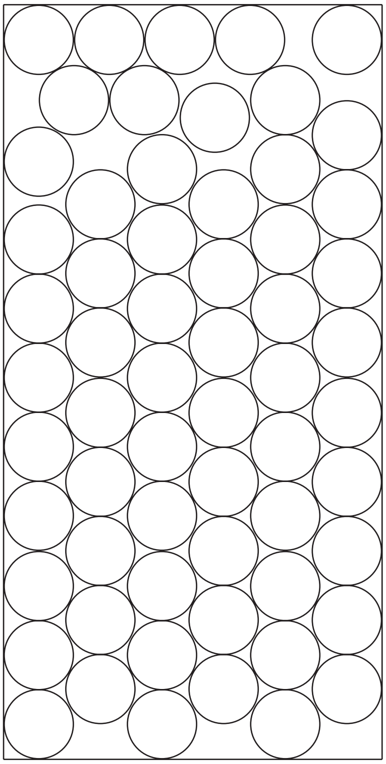


Figure 4: $R = 0.275$.

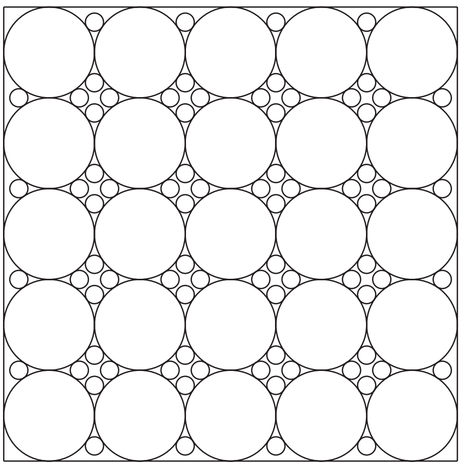


Figure 5

interrupted after the computational time exceeded 12-hour CPU time with *mipgap* = 8% achieved.

Figures 7 and 8 present packing pictures obtained by the formulation (13) for the quadratic container with $L = W = 60$ and four radii $R_1 = 0.7$, $R_2 = 2$, $R_3 = 4$, and $R_4 = 12$ with nesting permitted. Figure 7 corresponds to the grid with $M = N = 41$, while $M = N = 71$ was used for Figure 8. We see how packing approximation changes with the parameters of the grid. In Figure 7 the following circles were packed: 741 circles $R_1$, 52 $R_2$, 20 $R_3$, and 3 $R_4$. In Figure 8 we have 932 circles $R_1$, 77 $R_2$, 13 $R_3$, and 5 $R_4$ and there is still room for more circles. For the case of Figures 7 and 8 the

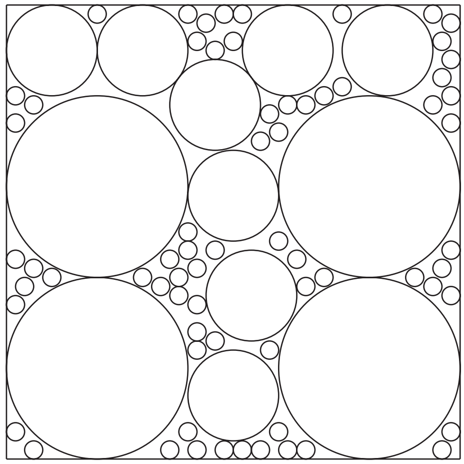


Figure 6

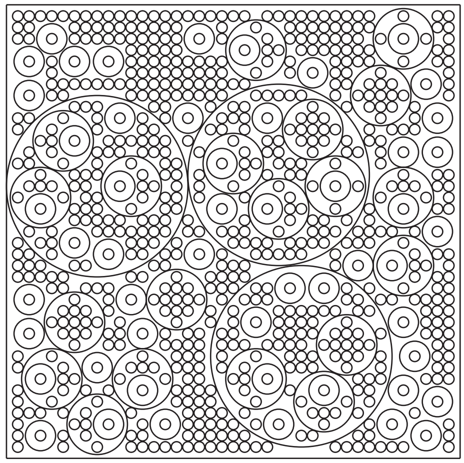


Figure 7

computations were interrupted after the computational time exceeded 12-hour CPU time with *mipgap* = 67% and *mipgap* = 83%, correspondingly.

## 4. Conclusions

Different integer formulations were proposed for approximate solution of the circle packing problem. We demonstrate that the pairwise formulation is stronger than the compact one obtained by summing up the nonoverlapping constraints. The presented approach can be easily generalized to the three- (and more) dimensional case and to different shapes of the container, including irregulars. We also proposed a formulation permitting nesting circles inside one another. This problem was mentioned in [7] in the context of packing pipes of different diameters into a shipping container and has not received much attention so far. An interesting direction for the future research is to study the use of Lagrangian relaxation and corresponding heuristics [3] to cope with large dimension of arising problems. Another area for future study is the use of valid inequalities for strengthening the formulations, as well as combining continuous and integer

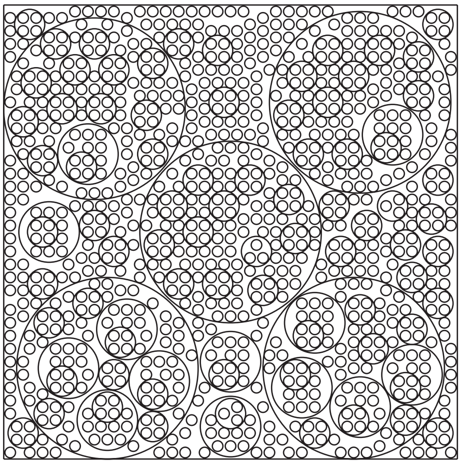


Figure 8

formulations in the solution process. Some complements in these directions are in course [16].

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was partially supported by Grants from RFBR, Russia (12_01_00893_a), and CONACYT, Mexico (167019). The authors would like to thank anonymous referees for their constructive comments and suggestions.

## References

[1] E. G. Birgin and J. M. Gentil, "New and improved results for packing identical unitary radius circles within triangles, rectangles and strips," *Computers and Operations Research*, vol. 37, no. 7, pp. 1318–1327, 2010.

[2] M. Hifi and R. M'Hallah, "A literature review on circle and sphere packing problems: models and methodologies," *Advances in Operations Research*, vol. 2009, Article ID 150624, 22 pages, 2009.

[3] I. Litvinchev, S. Rangel, and J. Saucedo, "A lagrangian bound for many-to-many assignment problems," *Journal of Combinatorial Optimization*, vol. 19, no. 3, pp. 241–257, 2010.

[4] E. Baltacioğlu, J. T. Moore, and R. R. Hill Jr., "The distributor's three-dimensional pallet-packing problem: a human intelligence-based heuristic approach," *International Journal of Operational Research*, vol. 1, no. 3, pp. 249–266, 2006.

[5] E. G. Birgin, J. M. Martínez, and D. P. Ronconi, "Optimizing the packing of cylinders into a rectangular container: a nonlinear approach," *European Journal of Operational Research*, vol. 160, no. 1, pp. 19–33, 2005.

[6] I. Castillo, F. J. Kampas, and J. D. Pintér, "Solving circle packing problems by global optimization: numerical results and industrial applications," *European Journal of Operational Research*, vol. 191, no. 3, pp. 786–802, 2008.

[7] H. J. Fraser and J. A. George, "Integrated container loading software for pulp and paper industry," *European Journal of Operational Research*, vol. 77, no. 3, pp. 466–474, 1994.

[8] J. A. George, "Multiple container packing: a case study of pipe packing," *Journal of the Operational Research Society*, vol. 47, no. 9, pp. 1098–1109, 1996.

[9] C. O. López and J. E. Beasley, "A heuristic for the circle packing problem with a variety of containers," *European Journal of Operational Research*, vol. 214, pp. 512–525, 2011.

[10] H. Akeb and M. Hifi, "Solving the circular open dimension problem using separate beams and look-ahead strategies," *Computers & Operations Research*, vol. 40, no. 5, pp. 1243–1255, 2013.

[11] M. H. Correia, J. F. Oliveira, and J. S. Ferreira, "Cylinder packing by simulated annealing," *Pesquisa Operacional*, vol. 20, pp. 269–286, 2000.

[12] C. O. Lopez and J. E. Beasley, "Packing unequal circles using formulation space search," *Computers & Operations Research*, vol. 40, pp. 1276–1288, 2013.

[13] Y. G. Stoyan and G. N. Yaskov, "Packing congruent spheres into a multi-connected polyhedral domain," *International Transactions in Operational Research*, vol. 20, no. 1, pp. 79–99, 2013.

[14] J. E. Beasley, "An exact two-dimensional non-guillotine cutting tree search procedure," *Operations Research*, vol. 33, no. 1, pp. 49–64, 1985.

[15] S. I. Galiev and M. S. Lisafina, "Linear models for the approximate solution of the problem of packing equal circles into a given domain," *European Journal of Operational Research*, vol. 230, no. 3, pp. 505–514, 2013.

[16] I. Litvinchev and E. L. Ozuna, "Packing circles in a rectangular container," in *Proceedings of the International Congress on Logistics and Supply Chain*, pp. 24–25, Queretaro, Mexico, October 2013.

[17] L. A. Wolsey, *Integer Programming*, Wiley, New York, NY, USA, 1999.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 695084, 9 pages
http://dx.doi.org/10.1155/2014/695084

*Research Article*

# US Natural Gas Market Classification Using Pooled Regression

**Vyacheslav V. Kalashnikov,[1,2,3] Gerardo A. Pérez-Valdés,[4] Timothy I. Matis,[5] and Nataliya I. Kalashnykova[3,6]**

[1] *Departamento de Ingeniería Industrial y de Sistemas, Tecnológico de Monterrey, Avenida Eugenio Garza Sada 2501 Sur, 64849 Monterrey, NL, Mexico*
[2] *Department of Experimental Economics, Central Economics and Mathematics Institute (CEMI), The Russian Academy of Sciences (RAS), Nakhimovsky Pr. 17, Moscow 117418, Russia*
[3] *Department of Electronics and Computing, Sumy State University, Rimsky-Korsakov Street 2, Sumy 40007, Ukraine*
[4] *Institutt for Industriell Økonomi og Teknologiledelse, Norges Teknisk Naturvitenskapelige Universitet, Alfred Getz Vie 3, 7491 Trondheim, Norway*
[5] *Department of Industrial Engineering, Texas Tech University (TTU), P.O. Box 43061, Lubbock, TX 79409, USA*
[6] *Departamento de Ciencias Físico-Matemáticas (FICA), Universidad Autónoma de Nuevo León (UANL), Avenida Universidad S/N, 66450 San Nicolás de los Garza, NL, Mexico*

Correspondence should be addressed to Vyacheslav V. Kalashnikov; kalash@itesm.mx

Received 25 December 2013; Accepted 18 February 2014; Published 23 March 2014

Academic Editor: Sergii V. Kavun



Natural gas marketing has considerably evolved since the early 1990s, when a set of liberalizing rules were passed in both the United States and the European Union that eliminated state-driven regulations in favor of open energy markets. These new rules changed many things in the business of energetics, and therefore new research opportunities arose. Econometric studies about natural gas emerged as an important area of study since natural gas may now be sold and traded in a number of stock markets, each one responding to potentially different behavioral drives. In this work, we present a method to differentiate sets of time series based on a regression model relating price, consumption, supply, and other factors. Our objective is to develop a method to classify different areas, regions, or states into groups or classes that share similar regression parameters. Once obtained, these groups may be used to make assumptions about corresponding natural gas prices in further studies.

## 1. Introduction

In the early 1990s, several regulations were passed in the US and the European Union [1–3] changing the way natural gas was marketed and traded. Particularly, this liberalization [4] effectively ended a period in which natural gas was a state-driven industry. The liberalization has also created the emergent natural gas markets, as well as a strong demand for models to better tackle the new problems and profit from this new setting [5, 6].

Owing not only to this liberalization, but also to the new local conditions that are more open to competition, new small players entered the natural gas industry, especially at the local scale. Indeed, the US has over 80 interstate, long-distance pipelines [7], serving different regions with various climatic, demographic, economic, and political circumstances. Natural gas usage in Alabama, for example, intuitively is not the same as in Oregon; thus the market dynamics of the fuel are also different, and this, we presume, should be reflected in some way in the econometric data of the states.

Not only macroeconomic trends, however, are affected by this setting. When doing cross-regions studies of various aspects of the supply chain, such as the forecasting of demand [8, 9], the balancing of the pipelines after imbalances have been created by the natural gas shippers [10–12], or the dynamics of interstate-intrastate systems [13], one has to take into account the existence of different markets. The existence of a common relationship between price and consumption of natural gas across several zones allows for strong claims of uniformity, which are useful when, for example, we are building scenarios for a stochastic problem. Indeed, if we manage to group the regions in clusters with similar price and consumption functions, we can reduce the number of variables needed in a scenario tree formulation [6, 14].

As such, we specify a regression function that relates many of the most relevant econometric figures for each of the 48 contiguous states of the American Union, modeling price as a function of explicative variables such as natural gas consumption, supply, and storage levels, as well as population (number of costumers), oil prices, temperatures, and production. The regression coefficients are then used to divide the set of states into several subsets or groups, obtaining a partition in which all the states in a group share the same regression parameters, and thus can be classified as an (implicit) market. The partition is made considering both statistical and nonstatistical characteristics of the obtained regression coefficients. The resulting partitions are next compared with others in their similitude and statistical significance, which would validate the goodness of the combination of the dendrogram and GRASP grouping methods.

This paper is organized as follows. The motivation and literature review on natural gas econometric regression is given in Section 2. Section 3 describes the way the regression function is derived, while Section 4 details the method for using the said function to perform the classification. Section 5 presents and discusses the results of the study, and conclusions are given in Section 6.

## 2. Natural Gas Price-Consumption Model

This work was motivated by our previous research in the natural gas supply chain, specifically developing an optimization model that addresses issues in interstate pipelines. The data used in this model, however, came from different regions, and therefore the time series involved did not necessarily behave in the same way.

As an example, suppose we are trying to model a certain problem that involves forecasting the residential consumption and price of natural gas in the states of Washington and Oregon, that is, four time series. If the robustness of the model is also a concern, then we should additionally consider different forecasting scenarios. Even with only two possible forecasting scenarios for each series (high/low consumption or prices) this translates into $2^4$ possible behaviors of the econometric parameters. If consumption is expressed as a function of price, however, then the scenario tree has only $2^2$ branches. Furthermore, if the regression function for both states is the same, then the number of scenarios can be reduced to just two. As the number of states being modeled scales up, that is, there are more than two parameters of interest, common assumptions like those mentioned above help reduce greatly the amount of scenarios in a stochastic model, optimization, or otherwise.

As we studied particular sets of data, it was noted that historical data of consumption and price showed conspicuous properties that could be used for the sake of our aims. Even though these data collections were taken from different states, all pairs of time series showed elastic consumption/demand [15, 16]; exponentially growing price averages [15, 17]; and both series in every pair seemed to be highly correlated to each other.

Indeed, the possibility of characterizing one set of series as a (regression) function of the other was interesting, as it would reduce the amount of data we needed to consider when modeling optimization problems. It is, of course, a common practice in economic and managerial sciences to do that since, for example, demand data is simpler to work with than price data [18]. The latter is mainly because the demand is usually easier to predict, and its behavior is less chaotic than that of prices. Such historical relationships between price and consumption is a common topic of study in time series economic analysis [19], which is mostly performed with the inclusion of other explicative variables, such as the price of substitutes (electricity, coal) and weather conditions.

This is the case of several models where the calculation of elasticities is the primary goal of the study [20]. Log-linear models [21–24] are generally favored because of the ease they provide when computing elasticity figures. However, linear models also have applications in the natural gas industry, like the Short-Term Integrated Forecasting system (STIFS) used by the United States Energy Information Agency in order to estimate natural gas demand as a function of several types of important variables related to the energy industry [25].

*2.1. Former and Current Approaches.* As explained in our previous work [26], a carefully designed regression function can help achieve such strong assumptions. Nevertheless, the study of such relationships and the possibility of forming state clusters based merely upon time series data analysis turned out to be interesting by itself, and we developed two different approaches to partition the collection of states. As we observed, neighboring states showed a large amount of diversity, yet different methods of grouping seemed to place certain states consistently together.

Two major areas of opportunity discovered were the design of the regression function, and the trade-off that each partition algorithm made use of.

Our previous paper [26] aimed at a very definite objective regarding the qualities of the regression model: it had to correlate consumption and price of residential natural gas series, using the former as the explicative variable because of the ease in its forecasting. The expression thus obtained served its purpose well, as demonstrated in its application to the optimization models in [27]; nevertheless, a more inclusive approach would involve series that comprise more information. Following the examples found in the literature and our own experience, we revealed that including more explicative series provided very good results in terms of regression fit. This has led to the model presented in the next section.

Coming back to the partitioning method, the two approaches presented before were as follows.

(i) The first one is the Dendrogram Grouping Method, which "cuts" a binary tree (whose nodes represent regression parameters) based on how close to each other the parameters are with respect to a given metric function and a weight scheme for the entries. This method proved replicative and fast, but it does not provide statistical significance to the grouped states' parameters (i.e., one state might find that temperature is a significant regressor, whereas some other state in the same group may not).

(ii) Another one is a greedy heuristic that starts with a number of states called "group leaders," and iteratively selects for each remaining state the group that suits the state best, based on its regression coefficient $R^2$. Because of the large amount of regressions performed, this method was reported to be slower and subject to accidental fluctuations, but the final results always guaranteed that all states in one group shared the same significance in their parameters.

In the following sections, we explain how we have improved over our latest approach, adding explicative power and robustness to the partitioning method and, ultimately, creating a better technique to identify similar regions based on their econometric data.

## 3. Regression Analysis

*3.1. Individual Multiple Linear Regression (IMLR).* Let $n$ be the total number of states, $m$ the number of observations per time series (months, in this case), $I = \{1, 2, \ldots, n\}$ the set of the 48 contiguous states of the American Union, $t \in \mathbf{T} = \{1, 2, \ldots, m\}$ the (discrete) time parameter, $\{P'_{i,t}\}$ the differenced residential natural gas price in state $i \in I$ at time $t \in \mathbf{T}$, $\{T'_{i,t}\}$ the differenced temperature, in Kelvin, shifted so that the minimum figure is $e$, $\{O'_t\}$ the differenced average spot price of oil in the US at time $t \in \mathbf{T}$, $\{N'_{i,t}\}$ the differenced number of residential consumers of natural gas in state $i \in I$ at time $t \in \mathbf{T}$, and $\{C'_{i,t}\}$ the differenced consumption of natural gas in state $i$ at time $t$.

Notice that all these series are *differenced*, or more precisely, lag-$(-1)$ differenced from the original values. This is because the said original values all tested positive for unit roots in the advanced Dickey-Fuller test. In contrast to the original series, the differenced series prove to be stationary; hence we make use of the latter.

This is the linear model we devised to relate the above-described series:

$$\widehat{C'_{i,t}} = \alpha_{0,i} + \alpha_{1,i}P'_{i,t} + \alpha_{2,i}C'_{i,t-12} + \alpha_{3,i}T'_{i,t} + \alpha_{4,i}O'_t + \alpha_{5,i}N'_{i,t}; \quad t \in \mathbf{T}; \quad i \in I. \tag{1}$$

We choose a Robust Regression Analysis using Huber weights to fit the series over traditional least-squares method due to nonnormality of the residuals experienced with the latter. Furthermore, due to the steps described in the next sections, heteroskedasticity would likely appear in the residuals once the pooling regression is carried on.

While most of the series were reasonably fit by (1), a couple of them showed very erratic behavior in either their natural gas price or consumption series. This is expected insofar economic forecasting is commonly subject to the large instability at time $t$. As the driving force behind short-term fluctuations in natural gas pricing is consumer demand rather than production supply, price was shown to be a significant factor when describing market consumption.

The selection of the descriptive variables was made considering other consumption models in the literature, the available data, and the significance found in the preliminary regression analysis. In particular, electricity prices and the natural gas supply and production, as well as a time index, were tested but found not to be significant in most of the states. This was especially interesting in the case of electricity prices, which certain sources cite as usual descriptors for the natural gas demand, but which were found to be 0.05 significant in only 12 of the 48 cases and thus dropped from the model.

The consumption and price of natural gas are endogenous variables as both are correlated to system shocks, such as unstable governments or weather-related events. As an alternative to the use of least squares regression to fit the model given in (1), a two-stage least squares approach could be employed with such instrumental variables as the number of gas producing wells, reserve estimates, and underground storage, to name only a few. However, this approach is not considered here, because the response (reaction) time of consumers' consumption habits to the shocks is much longer than that to the spot prices set by the market every day.

*3.2. Pooled Multiple Linear Regression (PMLR).* Now we address the issue of how one can use the same regression formula for more than one state, which would create several classes of states where demand responds to changes in the descriptors in a similar mode.

Assume that we have split $n$ collections of state time series into several classes, with the members of each class sharing a common set of regression parameters. Then the pooled data from the groups would be regressed at the same time, creating *pooled regressions*.

Let $\mathbf{I} = \{I_1, I_2, \ldots, I_K\}$ be a partition of $I$, and consider the model:

$$\widehat{C'_{i,t}} = \beta_{0,i} + \beta_{1,k}P'_{i,t} + \beta_{2,k}C'_{i,t-12} + \beta_{3,k}T'_{i,t} + \beta_{4,k}O'_t + \beta_{5,k}N'_{i,t}; \quad t \in T, \quad \forall i \in I_k, \quad k = 1, 2, \ldots, K. \tag{2}$$

Note that this model—called the Pooled Multiple Linear Regression (PMLR) model—has $K$ sets of parameters for each regressor variable, except for the intercepts $a_0^i$, which we allow to be different for each state. In comparison, model (1) has $n$ sets of parameters.

How should one define the partition $\mathbf{I}$ of the set of states? A good partition is expected to deliver groups of more or less congruent sizes, while maintaining a high individual $R^2$ value for each state. A good partition method should also be replicative (i.e., the same partition is obtained for the same group of states), be fast enough, and support the statistical significance.

## 4. Dendrogram-GRASP Grouping Method (DGGM)

In this section, a combination of both grouping methods mentioned in [26] into a GRASP heuristic is proposed. The resulting technique inherits the replicative property of the dendrogram method, while retaining the statistical significance of the heuristic algorithm.

*4.1. Dendrograms.* Dendrograms are binary trees in which two observation vectors $a$ and $b$ form the (sub-)branches of a higher branch $c$, so that

(i) these two observation vectors are "closer" to each other than to any other observation $d$,

(ii) $c$ is not an observation *per se*, but a new, artificial vector formed by some linear combination of $a$ and $b$.

The term "closer" is interpreted with respect to some metric (e.g., the Euclidean metric), while the artificial observations are produced by the weighted combination method. Once the dendrogram is formed, it is cut down from the root and thus generating (sub-) dendrograms with the branches resulting from the cut. The height of the cut is determined according to one of several criteria (the number of subdendrograms produced, the maximum allowed membership for the subdendrogram, etc.) The leaves pertaining to a given subdendrogram will pool their regression data together and form one group for the PMLR.

Previous experiments [26] have shown that what is called the "average euclidean" metric [28] delivers satisfactorily high $R^2$ levels with a better homogeneity in the resulting groups than other linkage function options.

*4.2. GRASP Heuristics.* GRASP stands for Greedy Randomized Adaptive Search Process; it is a metaheuristic, that is, a general method designed to provide good—but not necessarily optimal—results in problems otherwise too complicated to find an optimal solution, especially combinatorial problems [29].

Summarily, our GRASP approach will start with a seed formed by several one-state groups; then, for each state, it will identify those groups that deliver higher $R^2$ values once the data for the current state is pooled with that of the group. This is called the Restricted List of Candidates or RLC. A group $I_k$ from the RLC is chosen at random, and the current state is added to $I_k$, pooling its data with those already in the group. A number of swaps and movements are performed once the states are all in place, in order to try to improve the values of the resulting statistics $R^2$.

It is important to note that setting the values for the GRASP routine is rather subjective, since there is no definite objective to be achieved. Indeed, one cannot determine what number of groups is optimal, or which way is the best to define the greedy function. For example, one could prefer to increase the grouped $R^2$ value in each group rather than the average of the individual $R^2$s in that group, or vice versa. This is exemplified by the function

$$F_w(I_k) = \omega R^2_{I_k} + \frac{1-\omega}{|I_k|}\sum_{i\in I_k} R^2_i, \tag{3}$$

where

$$\begin{aligned} R^2_{I_k} &= 1 - \frac{\sum_{t\in\mathbf{T}, i\in I_k}\left(y_{i,t}-\widehat{y}_{i,t}\right)^2}{\sum_{t\in\mathbf{T}, i\in I_k}\left(y_{i,t}-\overline{y}_{I_k}\right)}, \\ R^2_i &= 1 - \frac{\sum_{t\in\mathbf{T}}\left(y_{i,t}-\widehat{y}_{i,t}\right)^2}{\sum_{t\in\mathbf{T}}\left(y_{i,t}-\overline{y}_{i}\right)}. \end{aligned} \tag{4}$$

Here, $y_{i,t} = \ln C'_{i,t}$, and $\overline{y}_\iota$ is understood as the average of all of the observations belonging to $\iota$ if the latter is a state (e.g., $\iota = i$) or as the average of the observations of the states in $\iota$, if the latter is a set of states (e.g., $\iota = I_k$).

For the local search, we handle the improvement function $G_\tau(I_k, I_\ell, I_i)$, which is used when deciding if it is convenient to move state $i$ from group $k$ to group $\ell$. It is parametrized by the improvement weight $\tau$:

$$G_\tau(I_k, I_l, I_i) = (1-\tau)\frac{R^2_{I_k}+R^2_{I_\ell}}{2} + \tau R^2_i. \tag{5}$$

*4.3. Dendrogram-GRASP Algorithm.* The following algorithm is used to classify the set of 48 contiguous states of the United States into groups that share a common regression function.

(1) Initialize the values for each of the time series in each of the 48 states. Set a seed size $s_{\text{Seed}}$, a maximum number of groups $s_{\text{Max}}$, a RLC size $s_{\text{RLC}}$, an individual/grouped $R^2$ weight $\omega \in [0,1]$, an individual/grouped threshold $\varphi \in (0,1)$, an improvement weight $\tau \in (0,1)$, a relative improvement threshold $\psi \in [0,1]$, and a maximum number of local search steps, $s_{\text{ls}}$.

*Seed Selection*

(2) Perform an IMLR on each of the 48 sets of time series, obtaining $\alpha_{j,i}$, $j = 0,\ldots,6$, $i \in I$.

(3) Form a dendrogram of 48 leaves with the vectors $\alpha$, using the average euclidean mean as the linkage function, and cut it so that there are exactly $s_{\text{Seed}}$ subtrees.

(4) Select the state with the highest $R^2_i$ from each of the obtained groups and call it the $k$th group's leader. Define the one-state groups obtained as the partition $\mathbf{I}_k$. All the nonselected (spare) states form the set *Active*.

*Greedy Process*

(5) For each state $x$ in the set *Active*,

(a) pool the data of $x$ with the data of each of the formed groups and perform a pooled regression; select a number of $s_{\text{RLC}}$ groups with the highest value of the greedy function $F_w$ and form the RLC;

(b) choose randomly one of the groups from the RLC, for example, $I_a$.

(i) If none of the candidate groups in the RLC delivers $F_w(I_k) > \varphi$ and we have not yet reached the maximum number of groups $s_{\text{Max}}$, create a new group $I_x = \{x\}$ containing only $x$, remove $x$ from the active set, and update all the parameters.

(ii) Otherwise, assign $x$ to $I_a$, remove $x$ from the active set, and update all the parameters.

(6) All of the states are now partitioned into the groups, and we can begin the local search.

*Local Search*

(7) For $l = i$ to $l = s_{\text{ls}}$,

(a) randomly select one of the formed groups, $I_a$, and one state in that group, $x$; select another group, $I_b$; compute $g_1 = G_\tau(I_a, I_b, x)$;

(b) remove $x$'s data from $I_a$ and pool the same data of $x$ with $I_b$; compute $g_2 = G_\tau(I_a, I_b, x)$;

(c) if $g_1 \geq (1 + \psi)g_2$, remove $x$ from $I_b$ and return it to $I_a$; otherwise, continue.

(8) Report the obtained groups as the desired partition.

(9) End.

*4.4. Partition Similarity.* To determine the similitude of two partitions, we will use an expression that, roughly speaking, counts the number of coincidences found in two partitions and divides it by the number of total possible coincidences, given the sizes of the groups in each partition. While there are many disputable ways to measure the similitude between partitions with a different number of elements, this method was chosen because of its normality. Indeed, it will always return 1 when both partitions are identical and will always return 0 when there are no coincidences between two partitions, that is, when no two states share a group in both partitions and no state is single-grouped in both partitions.

Let $\mathbf{I} = \{I_1, I_2, \ldots, I_K\}$, $\mathbf{J} = \{J_1, J_2, \ldots, J_L\}$ be two arbitrary partitions of the set of states, with $I_i = \{I_1^i, I_2^i, \ldots, I_{k^i}^i\}$, $i = 1, \ldots, K$, and $J_j = \{J_1^j, J_2^j, \ldots, J_{l^j}^j\}$, $j = 1, \ldots, L$.

The function $\mathbf{a}_{\mathbf{I},\mathbf{J}}$ defined by

$$\mathbf{a}_{\mathbf{I},\mathbf{J}}(I_i) = \begin{cases} 1, & \text{if } I_i = \{m\} = J \text{ for any } J \in \mathbf{J}, \\ 0, & \text{otherwise,} \end{cases} \qquad I_i \in \mathbf{I}, \quad (6)$$

assumes the value 1 if group $I_i$ contains a single state in partition $\mathbf{I}$ and this state also forms a group-singleton in partition $\mathbf{J}$.

For every pair of states, we will assess if they share a group in a given partition using the following function $\mathbf{b}_{\mathbf{J}}$:

$$\mathbf{b}_{\mathbf{J}}(m, n) = \begin{cases} 1, & \text{if } m, n \in J_j, \text{ for any } j; \\ 0, & \text{otherwise,} \end{cases} \qquad m, n \in \mathbf{I}. \quad (7)$$

In case the function $\mathbf{a}_{\mathbf{I},\mathbf{J}}$ has the value of 1, we say that we have a (one-state) coincidence, which means that the state has been found incompatible with other states twice, no matter which method formed partitions $\mathbf{I}, \mathbf{J}$.

Similarly, if the function $\mathbf{b}_{\mathbf{J}}$ returns 1 for two states *in a group from the partition* $\mathbf{I}$, we say that we have a (two-state) coincidence; that is, in both partitions, the two states are members of the same group.

To measure the number of coincidences between two partitions, we use the function:

$$\begin{aligned} C_q(I_i, \mathbf{I}, \mathbf{J}) &= \mathbf{a}_{\mathbf{I},\mathbf{J}}(I_i) + (1 - \mathbf{a}_{\mathbf{I},\mathbf{J}}(I_i)) \\ &\quad \times \left( \sum_{m \in I_i} \sum_{n \in I_i, n \neq m} \frac{q\mathbf{b}_{\mathbf{J}}(m, n) + (1 - q)}{2} \right), \end{aligned} \quad (8)$$

for $I_i \in \mathbf{I}$, $q = \{0, 1\}$.

If the parameter $q$ equals 1, then the function $C_q$ counts the number of either type of coincidences that couples of states reveal in the group $I_i$ in comparison to the groups they belong to in the partition $\mathbf{J}$. Conversely, if $q = 0$, then we simply count the total number of possible coincidences for the states in the group $I_i \in \mathbf{I}$. Note that the function $C_q$ is not necessarily symmetric with respect to the pairs of partitions: $C_q(I_i, \mathbf{I}, \mathbf{J})$ need not have the same value as $C_q(I_i, \mathbf{J}, \mathbf{I})$.

The similitude function used Sim is defined as follows:

$$\text{Sim}(\mathbf{I}, \mathbf{J}) = \frac{\sum_{I_i \in \mathbf{I}} C_1(I_i, \mathbf{I}, \mathbf{J}) + \sum_{J_j \in \mathbf{J}} C_1(J_j, \mathbf{J}, \mathbf{I})}{\sum_{I_i \in \mathbf{I}} C_0(I_i, \mathbf{I}, \mathbf{J}) + \sum_{J_j \in \mathbf{J}} C_0(J_j, \mathbf{J}, \mathbf{I})}. \quad (9)$$

Notice that if there is at least one group in either partition containing more than one element, then $C_0$ for that group is at least 1, whereas if there exists no such group in either partition, then $\mathbf{a}_{\mathbf{I},\mathbf{J}}(a) = 1$ and consequently $C_0(a, \mathbf{I}, \mathbf{J}) = 1$ for any $a \in \mathbf{I} \cap \mathbf{J}$. Therefore, the denominator is never 0, which makes this function well defined.

**Lemma 1.** *Let* $\mathbf{I}$ *and* $\mathbf{J}$ *be two partitions of the set* $I = \{1, 2, \ldots, n\}$, *and let function* Sim *be defined by* (9). *The following statements are true:*

(1) $\text{Sim}(\mathbf{I}, \mathbf{J}) = \text{Sim}(\mathbf{J}, \mathbf{I})$;

(2) $\text{Sim}(\mathbf{I}, \mathbf{J}) = 1$ *if and only if* $\mathbf{I} = \mathbf{J}$;

(3) $\text{Sim}(\mathbf{I}, \mathbf{J}) = 0$ *if and only if there are neither one-state nor two-state coincidences between* $\mathbf{I}$ *and* $\mathbf{J}$;

(4) $0 \leq \text{Sim}(\mathbf{I}, \mathbf{J}) \leq 1$.

*Proof.* (1) This is easy to see from the structure of the function.

(2) Let $\mathbf{I} = \mathbf{J}$. If $I_i = \{m\} = J_k$ for some $i$ and $k$, then $C_1(I_i, \mathbf{I}, \mathbf{J}) = C_0(I_i, \mathbf{I}, \mathbf{J})$. Otherwise, if the order of $I_i$ is greater than one, then the second term in (8) (the definition of $C_q$) assumes the same value no matter whether $q = 1$ or $q = 0$. Therefore, the numerator and denominator in Sim are equal.

Conversely, if there exists one $I_i$ such that $I_i \neq J$ for all $J \in \mathbf{J}$, then $C_1(I_i, \mathbf{I}, \mathbf{J})$ is strictly less than $C_0(I_i, \mathbf{I}, \mathbf{J})$. Since $C_1(J_i, \mathbf{I}, \mathbf{J}) \leq C_0(J_i, \mathbf{I}, \mathbf{J})$, it follows that the numerator in (9) (defining Sim) is strictly smaller than the denominator, and therefore $\mathrm{Sim}(\mathbf{I}, \mathbf{J}) < 1$.

(3) If there is at least one one-state coincidence, or a two-state coincidence, then the numerator in Sim is larger than 0, and therefore $\mathrm{Sim}(\mathbf{I}, \mathbf{J}) > 0$.

Conversely, since $C_q$ is nonnegative for every value of $q$, $\mathrm{Sim}(\mathbf{I}, \mathbf{J}) = 0$ means that both terms in the numerator are zero, which is only possible if $\mathbf{a}_{\mathbf{I},\mathbf{J}}(I_i) = \mathbf{a}_{\mathbf{I},\mathbf{J}}(J_j) = 0$ for every member of $\mathbf{I}$ and $\mathbf{J}$, *and* $\mathbf{b}_{\mathbf{J}}(m, n) = 0$ for every $m, n \in I$, which means that there is no coincidence of any type between these two partitions.

(4) The first inequality follows from the fact that both the numerator and denominator in (9) are positive. The second inequality comes from the same argument as in item (2); that is, the numerator is either equal or strictly less than the denominator. □

## 5. Experimental Results

This section presents the results of the numerical experimentation performed on a number of times series pertaining to each of the 48 data sets. The values for the historical natural gas prices, consumption, and number of consumers, as well as the oil spot prices were taken from the US Energy Information Agency, whereas the temperature figures for each state were obtained from the US Department of Commerce National Oceanographic and Atmospheric Agency [30].

*5.1. IMLR Results.* The first step was to perform the IMLR for the 48 sets of time series; this provided the regression parameters for the dendrogram formation. The five time series corresponding to every state had 240 monthly observations each.

Individual regression models showed regression $R^2$ coefficients with the average of 0.77 and the minimum of 0.61. The normality and heteroskedasticity were not tested due to the use of Robust Regression with Huber weights. Randomness of the residuals was tested, and high $P$ values were found for many states.

*5.2. Dendrogram-GRASP Grouping Results.* There are two main aspects we wanted to consider when evaluating the effectiveness of the Dendrogram-GRASP approach: how replicative it is, and how good a partition is produced. The first issue is evaluated by examining how good and how similar the partitions are that come from the same seed (as opposed to those that come from randomly generated seeds). The goodness of one partition is measured with the average group [state] coefficient of determination, $R^2_{I_k}$ [$R^2_i$], calculated across all the groups [states] of the partitions.

There are, however, a number of different design parameters that should be included in the experimentation. Each experimental observation consists of the generation of 10 partitions, using the following parameters.

(i) A seed choice: the dendrogram seed (DDR), a random seed common to all 20 partitions (FIX), and a random seed for each partition (RND).

(ii) The individual versus grouped $R^2$ weight, $\omega$, which determines what is more important when adding a new state to an existing group in the GRASP routine: values considered in the experimentation are $\omega = 0$ (only the single states' $R^2$s are considered), 0.5, and 1 (only the groups' $R^2$s are important).

(iii) The new group threshold, $\varphi$: the closer the value of $\varphi$ to 1, the more likely new single-state groups will be created in the GRASP routine. The tested values are $\varphi \in \{0.90, 0.95\}$.

(iv) The length of the restricted candidate list, $s_{\mathrm{RCL}}$: the values considered are $s_{\mathrm{RLC}} \in \{1, 5\}$.

(v) The number of local search moves: $s_{\mathrm{ls}} \in \{0, 100\}$.

(vi) The local search individual/grouped $R^2$ weight, $\tau$: considered values are $\tau \in \{0, 0.66, 1\}$.

The starting number of groups was fixed at 10, and the maximum number of groups allowed was set at 15. Each combination of levels was replicated 20 times. This resulted in 5760 experimental observations.

In each observation, we calculated the average similitude between the various partitions involved, as well as their similitude with a randomly created partition. The compared similitudes were as follows:

(i) the average similitude of the dendrogram partition to each of the 20 GRASP partitions (DG);

(ii) the average similitude of a random partition and each of the 20 GRASP partitions (GR);

(iii) the average similitude of the 20 GRASP partitions among themselves (GG).

The first part of the analysis consisted in testing all the experimental observations. After that, only the most convenient levels were kept.

Tables 1 and 2 present a summary of the results of the experimental runs. The first three data columns show the average similarities for each of the three comparisons of interest, whereas the last two columns show the average of the individual and grouped coefficients of determination.

A quick look at this table suggests that the similitude figures are characteristically low: the average similarity of an arbitrary partition to a randomly formed one, calculated using all the observations, is 0.0947. This will be called the partitions' randomness. If columns 3 and (particularly) 5 approach the average randomness for this experiment, the partition method is not very efficient. This especially concerns the cases $s_{\mathrm{ls}} = 5$, $\omega = 0$, and $\tau = 1$, whose similarity measures are fairly low. Luckily enough, in all these cases the average GG similarities were found to be statistically different

Table 1: Experimental Results I.

| Factor | Level | Av. similitude | | | Av. $R^2$ values | |
|---|---|---|---|---|---|---|
| | | DG | GR | GG | Av. $R^2_{I_i}$ | Av. $R^2_{I_k}$ |
| $\varphi$ | 0.90 | 0.145 | 0.079 | 0.178 | 0.503 | 0.535 |
| | 0.95 | 0.149 | 0.079 | 0.177 | 0.499 | 0.537 |
| Seed | DDR | 0.182 | 0.083 | 0.194 | 0.513 | 0.568 |
| | FIX | 0.130 | 0.077 | 0.154 | 0.489 | 0.521 |
| | RND | 0.128 | 0.077 | 0.184 | 0.501 | 0.520 |
| $\omega$ | 0 | 0.146 | 0.082 | 0.136 | 0.427 | 0.564 |
| | 0.5 | 0.147 | 0.079 | 0.183 | 0.534 | 0.535 |
| | 1 | 0.148 | 0.077 | 0.213 | 0.542 | 0.511 |
| $\tau$ | 0 | 0.160 | 0.080 | 0.232 | 0.699 | 0.502 |
| | 0.66 | 0.141 | 0.079 | 0.150 | 0.455 | 0.554 |
| | 1 | 0.140 | 0.079 | 0.149 | 0.349 | 0.553 |

Table 2: Experimental Results II.

| Factor | Level | Av. similitude | | | Av. $R^2$ values | |
|---|---|---|---|---|---|---|
| | | DG | GR | GG | Max. $R^2_{I_i}$ | Max. $R^2_{I_k}$ |
| $s_{\mathrm{RLC}}$ | 1 | 0.160 | 0.082 | 0.247 | 0.879 | 0.862 |
| | 5 | 0.134 | 0.076 | 0.107 | 0.879 | 0.871 |
| $s_{\mathrm{ls}}$ | 0 | 0.167 | 0.084 | 0.236 | 0.876 | 0.857 |
| | 100 | 0.126 | 0.075 | 0.119 | 0.882 | 0.876 |

(higher) than their respective GR similarities by making use of the Wilcoxon signed-rank (WSR) $\alpha = 0.95$ test.

The average $R^2$ values in columns 6 and 7 do not deviate much from the averages across all the observations, 0.602 and 0.624, respectively, with the exception of the grouped individual parameter $R^2_{I_i}$ for $\tau = 1$. It is clear that certain similarity values for some levels are consistently lower than others. There is, for example, a very large difference between the average DG similitude obtained using a DDR seed than using a RND or FIX seed and so on. Based on this, we decided to discard some of the levels whose averages are not only considerably lower, but also the observations for each level are determined to be different by a WSR test.

Now let us look at each of the level values we should consider to drop. The first level, the GRASP new group threshold $\varphi$, shows a very similar GG figure, and equally similar $R^2$ values. We decide to keep the factor levels intact, in case these figures change once other levels are removed.

Seeds are more difficult to assess. The FIX seed shows lower values than the DDG one, but still higher than the RND. Weight $\tau$ shows much better numbers in all but the grouped $R^2$ entry. Because of this, we pick it as the only label for the later study. On the contrary, $\omega$ is better at value 1, except again in the grouped $R^2$ column. This result for $\omega$ is very counter intuitive! However, the two values serve a similar purpose at different parts of the process, so this behavior might indeed be justified.

The factors $s_{\mathrm{RLC}}$ and $s_{\mathrm{ls}}$ were introduced to add variation in the GRASP routine, and their results appear separated in Table 2. This is because, while their similitude values work in the same way as the other factors, the $R^2$ measurements per observation are not the average across all 10 partitions in the observation, but rather the maximum obtained. In a common GRASP routine, the process will be repeated several times and the best solution will be adopted. For our case, this means that we should choose the best of the 20 partitions in each observation, and this decision becomes the result for that observation. Arguably, both the individual and grouped average maximum coefficients of determination seem to show little difference. In particular, the differences are deemed not large enough to justify the trade-off with similarity in all cases. While this was expected from the extended RLC size, the poor results obtained by the local search suggest that we should rethink our local search procedure in the future.

Based on similarity alone, we decided to eliminate the poorest levels and kept only a single-group state list and a zero-swaps local search for the second part of the analysis. After deciding to drop several levels, we will rewrite the results table including only the accepted levels, to see how the figures change once the poorest results are winnowed.

The much smaller Table 3 is the consequence of fixing $\omega = 1$, $\tau = 0$, $s_{\mathrm{RLC}} = 1$, and $s_{\mathrm{ls}} = 0$ and eliminating the RND seed choice, which results in 100 observations. Now the similitudes look much better: we have the sample average of 0.438 and the maximum of 0.477, which means that, for the parameters chosen, the similitudes obtained are remarkably higher than the average randomness.

For the first factor, $\varphi$, the similitudes are of little difference, the same as the determination coefficients in all accounts. However, for the seed levels, the DDR seed clearly favors similitude between the seed and the resulting partition.

Table 3: Experimental Results III.

| Factor | Level | Av. similitude | | | Av. $R^2$ values | |
|---|---|---|---|---|---|---|
| | | DG | GR | GG | Av. $R_i^2$ | Av. $R_{I_k}^2$ |
| $\varphi$ | 0.90 | 0.171 | 0.077 | 0.432 | 0.760 | 0.340 |
| | 0.95 | 0.178 | 0.085 | 0.432 | 0.759 | 0.349 |
| Seed | DDR | 0.238 | 0.090 | 0.454 | 0.757 | 0.432 |
| | FIX | 0.143 | 0.074 | 0.365 | 0.760 | 0.299 |
| | RND | 0.143 | 0.079 | 0.477 | 0.761 | 0.302 |

Similitude among resulting partitions is also good at the RND partition, which could indicate the particular FIX seed was initially a bad choice when compared to either an average partition seed or one selected in a methodical way.

The coefficients of determination $R^2$ present a rather interesting development. The individual coefficients $R_i^2$ are decent enough when compared to the ones from the dropped levels, but there is a dramatic drop in the group figures $R_{I_k}^2$, which decreased from an average of around 0.53 to as low as 0.299. This happens because, while focusing on similitude, we chose in favor of $s_{\text{ls}} = 0$, which yields the mean $R_{I_k}^2$ of only 0.366, as opposed to the 0.706 value obtained after fixing $s_{\text{ls}} = 100$. In Table 2, however, we see the greater max $R_{I_k}^2$ because it was relevant to that table. If we were to remake Table 3 using the value of $s_{\text{ls}} = 100$ for this level, similitudes would fall around 10%, but the average group determination coefficients $R_{I_k}^2$ would increase to roughly 0.43, which is much better than that with $s_{\text{ls}} = 0$. Maximum values for the different $R^2$s, correspondent to those in Table 3, remain mostly unchanged.

## 6. Concluding Remarks

In this paper, we propose and justify a heuristic method to group several zones based on a regression function that estimates several factors related to the natural gas demand. The groups thus obtained share key information regarding the behavior of natural gas-related historic econometric data.

We start by developing a linear regression model that correlates natural gas historic residential consumption and several explicative variables, such as the residential price, number of consumers, and temperature. This model, inspired by several examples in the literature, fits well the time series employed and has good predictive power, but it is by no means the only one that can be used nor necessarily the best.

The results of each of the 48 regressions performed are then used to create dendrogram-based partitions, which are in turn used as the starting point in a GRASP routine. The latter, while tending to form rather dissimilar partitions (compared to the dendrogram grouping), has the advantage of adding statistical significance to all the regressions in all the groups formed.

We tested several parameters in an experimental design consisting of more than 4300 observations, six factors, and two or three levels per factor. Using ad hoc and nonparametric selections, we tried to obtain a good combination of parameters, namely, one that delivers high similitude between partitions obtained from the same seed and a satisfactory goodness of the pooled regressions.

Similitude is measured by a standardized function which equals 0 if there are no common groups between two partitions of a fixed set and 1 if both partitions are identical. We were able to obtain experimental conditions with similitudes (mostly) above 0.43, which are deemed good considering that the average randomness of a partition in the study is around 0.09.

It is encouraging that, using the regression function herein proposed, the GRASP routine worked well by itself and also when combined with the dendrogram partitioning method. Unfortunately, the inclusion of randomness did not provide for good results, as it offered no increase in goodness of the partitions but a considerable decrease in similitude when a long RLC was used. The proposed local search approach was found to have a negative impact on the similitude values, though not overly so. However, at the same time it did affect heavily the values of the grouped coefficients of determination when the maximum values were considered in the selection but the averaged values were looked into in the end results. The "goodness" of the regressions, as discussed, must then be judged with a more nuanced approach.

The entire work frame summarized here is intended to provide a way to identify individuals (states, in this case) with common econometric behavior among themselves by means of statistically significant information. Such results used to help us in the past in the context of optimization theory (by greatly decreasing the number of variables in stochastic problems), and we believe this technique has other applications in economic analysis.

The planned future work includes enhancing the robustness of the method by designing better GRASP RLC and local search procedures, trying sampled regressions when forming large groups to gain on time and studying how different data sets and regression models would work in combination with the Dendrogram-GRASP approach proposed in the paper.

## Conflict of Interests

The authors declare that there is no conflict of interests for any of the authors of the paper.

## Acknowledgments

The research activity of the first author was financially supported by the R&D Department (Cátedra de Investigación)

CAT-174 of the Tecnológico de Monterrey (ITESM), Campus Monterrey and by the SEP-CONACYT Project CB-2008-01-106664, Mexico. Also, the work of the fourth author was supported by the National Council of Science and Technology (CONACyT) of Mexico as part of the Project CB-2011-01-169765, PROMEP 103.5/11/4330, and PAICYT 464-10.

## References

[1] Energy Information Administration, "FERC Order 636: The Restructuring Rule," 2005, http://www.eia.doe.gov/emeu/steo/pub/document/textng.html.

[2] Energy Information Administration, "FERC Policy on System Ownership Since 1992," 2005, http://www.eia.doe.gov/oil_gas/natural_gas/analysis_publications/ngmajorleg/fercpolicy.html.

[3] A. Soto, "FERC Order 636 & 637," 2008, http://www.aga.org/Legislative/issuesummaries/FERCOrder636637.html.

[4] IHS Engineering, "EC Proposes New Legislation for European Energy Policy," 2007, http://engineers.ihs.com/news/eu-en-energy-policy-9-07.html.

[5] Environmental Protection Agency, "The Impacts of FERC Order 636 on coal mine gas project development," 2008, http://www.epa.gov/cmop/docs/pol004.pdf.

[6] K. T. Midthun, *Optimization Models for Liberalized Natural Gas Markets*, Norwegian University of Science and Technology. Faculty of Social Science and Technology Management Department of Industrial Economics and Technology Management, Trondheim, Norway, 2009.

[7] M. J. Doane and D. F. Spulber, "Open access and the evolution of the US spot market for natural gas," *Journal of Law and Economics*, vol. 34, no. 2, pp. 447–517, 1994.

[8] R. Gutiérrez, A. Nafidi, and R. G. Sánchez, "Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model," *Applied Energy*, vol. 80, no. 2, pp. 115–124, 2005.

[9] F. K. Lyness, "Gas demand forecasting," *Journal of the Royal Statistical Society—Series D*, vol. 33, no. 1, pp. 9–12, 1984.

[10] S. Dempe, V. Kalashnikov, and R. Z. Ríos-Mercado, "Discrete bilevel programming: application to a natural gas cash-out problem," *European Journal of Operational Research*, vol. 166, no. 2, pp. 469–488, 2005.

[11] V. V. Kalashnikov and R. Z. Ríos-Mercado, "A natural gas cash-out problem: a bilevel programming framework and a penalty function method," *Optimization and Engineering*, vol. 7, no. 4, pp. 403–420, 2006.

[12] N. Keyaerts, L. Meeus, and W. D'haeseleer, "Analysis of balancing-system design and contracting behavior in the natural gas markets," in *Proceedings of the European Doctoral Seminar on Natural Gas Research*, Delft, The Netherlands, 2008.

[13] H. G. Huntington, "Federal price regulation and the supply of natural gas in a segmented field market," *Land Economics*, vol. 54, no. 3, pp. 337–347, 1978.

[14] A. Tomasgard, F. Romo, M. Fodstad, and K. T. Midthun, "Optimization models for the natural gas value chain," in *Geometric Modeling, Numerical Simulation, and Optimization: Applied Mathematics at SINTEF*, optimization models for the natural gas value chain, Springer, 2007.

[15] C. Nelder, "Natural Gas Price ForecastThe future of Natural gas: It's time to invest," 2009, http://www.energyandcapital.com/articles/natural-gas-price-forecast/916.

[16] EuroGas, "EuroGaslong term outlook to 2030," 2009, http://www.eurogas.org/uploads/media/Statistics_Eurogas_LT_Outlook_2007-2030_Final_25.11.10.pdf.

[17] P. W. MacAvoy, *The Natural Gas Market: Sixty Years of Regulation and Deregulation*, Yale University Press, New Haven, Conn, USA, 2000.

[18] K. T. Talluri and G. J. van Ryzin, *The Theory and Practice of Revenue Management*, Springer, New York, NY, USA, 2004.

[19] P. W. Keat and P. K. Y. Young, *Managerial Economics: Economic Tools for Today's Decision Makers*, Prentice Hall, Englewood Cliffs, NJ, USA, 2006.

[20] J. M. Gowdy, "Industrial demand for natural gas. Inter-industry variation in New York state," *Energy Economics*, vol. 5, no. 3, pp. 171–177, 1983.

[21] J. G. Beierlein, J. W. Dunn, and J. G. McConnon Jr., "The demand for electricity and natural gas in the northeastern United States," *The Review of Economics and Statistics*, vol. 63, no. 3, pp. 403–408, 1981.

[22] W. T. Lin, Y. H. Chen, and R. Chatov, "The demand for natural gas, electricity and heating oil in the United States," *Resources and Energy*, vol. 9, no. 3, pp. 233–258, 1987.

[23] N. Krichene, "World crude oil and natural gas: a demand and supply model," *Energy Economics*, vol. 24, no. 6, pp. 557–576, 2002.

[24] S.-H. Yoo, H.-J. Lim, and S.-J. Kwak, "Estimating the residential demand function for natural gas in Seoul with correction for sample selection bias," *Applied Energy*, vol. 86, no. 4, pp. 460–465, 2009.

[25] Energy Information Administration, "Natural Gas Model Description," 1999, http://www.eia.doe.gov/oil_gas/natural_gas/analysis_publications/ngmajorleg/ferc636.html.

[26] V. V. Kalashnikov, T. I. Matis, and G. A. Pérez-Valdés, "Time series analysis applied to construct US natural gas price functions for groups of states," *Energy Economics*, vol. 32, no. 4, pp. 887–900, 2010.

[27] V. V. Kalashnikov, G. A. Pérez-Valdés, A. Tomasgard, and N. I. Kalashnykova, "Natural gas cash-out problem: bilevel stochastic optimization approach," *European Journal of Operational Research*, vol. 206, no. 1, pp. 18–33, 2010.

[28] Mathworks Inc, "Statistics Toolbox: Linkage," 2008, http://www.mathworks.com.

[29] P. Festa and M. G. C. Resende, "GRASP: an annotated bibliography," in *Essays and Surveys in Metaheuristics*, pp. 325–367, Citeseer, 2002.

[30] "NOAA Satellite and Information Service," 2010, ftp://ftp.ncdc.noaa.gov/pub/data/cirs/.

Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2014, Article ID 816286, 13 pages
http://dx.doi.org/10.1155/2014/816286

*Research Article*

# Variations in the Flow Approach to CFCLP-TC for Multiobjective Supply Chain Design

**Minor P. Hertwin,[1] Olivares-Benítez Elías,[2] Tapia-Olvera Ruben,[1] and Martínez Flores J. Luis[2]**

[1] *School of Engineering, Polytechnic University of Tulancingo, Tulancingo 43629, Mexico*
[2] *School of Engineering, UPAEP University, Puebla 72410, Mexico*

Correspondence should be addressed to Minor P. Hertwin; hertwin.minor@upt.edu.mx

Received 18 October 2013; Accepted 25 December 2013; Published 13 March 2014

Academic Editor: Timothy I. Matis



We review the problem for the design of supply chains called Capacitated Fixed Cost Facility Location Problem with Transportation Choices (CFCLP-TC). The problem is based on a production network of two echelons with multiple plants, a set of potential distribution centers, and customers. The problem is formulated as an optimization model with two objective functions based on time and cost. This paper proposes three changes to the original model to compare the sets of efficient solutions and the computational time required to obtain them. The main contribution of this paper is to extend the existing literature by incorporating approaches for the supply of product to customers through multiple sources, the direct flow between plants and customers, without this necessarily implying removing the distribution centers, and the product flow between distribution centers. From these approaches, we generate mathematical programming models and propose to solve through the epsilon-constraint approach for generating Pareto fronts and thus compare each of these approaches with the original model. The models are implemented in GAMS and solved with CPLEX.

## 1. Introduction

Supply Chain Management (SCM) is the process of planning, implementing, and controlling the operation of the supply chain efficiently. SCM spans all movements and storage of raw materials, work-in-process inventory, and finished goods from the point of origin to the point of consumption [1]. Part of the planning processes in SCM aim at finding the best possible supply chain configuration so that all operations can be performed in an efficient way. The Capacitated Fixed Cost Facility Location Problem with Transportation Choices (CFCLP-TC) proposed by [2, 3] is a combinatorial optimization problem for supply chain design. It is an extension of the CFLP as a biobjective mixed-integer program. It is based on a two-echelon system for the distribution of one product in a single time period with two objectives: to minimize cost and to minimize the time of transportation from plants to customers. This approach considers several alternatives to transport the product from one facility to the other in each echelon of the network. The criterion of cost is an aggregate function of variable cost and fixed cost. At difference from similar works in the literature, the aim here is to provide the decision maker with a set of nondominated alternatives to allow her deciding. Some qualitative information only known by the decision maker may motivate the selection of one of those alternatives.

This paper presents three innovative approaches for the design of a supply chain for biobjective problems. In the first approach, it is proposed that customers can be supplied by more than one distribution center; in the second approach it is proposed that customers can be supplied directly by plants, but the distribution centers can be used. In the third approach proposed here a product flow may exist between the distribution centers. For the study of this problem and the proposed variations, instances of different sizes were used. The results are compared on the basis of three metrics. The first metric is called $R_{\text{pos}}$, proposed by [4], and the second and third metrics are called $D_{\text{avg}}$ and

$D_{\min}$, proposed by [3] to compare two biobjective Pareto fronts.

This paper is divided into five sections as follows: the first section presents a general introduction to the work. The second section presents the literature review to define the opportunity area that the work will fill. The third section presents an overview of the problem and its variations. The fourth section describes the computational experiment, the metrics used to evaluate the Pareto fronts, and the results of the computational implementation. Finally, in the fifth section the conclusions and proposals to carry out further work are exposed.

## 2. Literature Review

Historically, researchers have focused on the design of distribution systems [5], without considering the supply chain as a whole. Typically, discrete location models were proposed to include additional features. Reference [6] reviews some important mixed-integer formulations for productions-distribution systems. However, those models had limited scope and could not cope with realistic supply chain structures. Reference [5] proposes the inclusion of relevant features for the Supply Chain Management (SCM) in the facility location models that gradually began to be considered. These include

(a) subsets of the products (customer specifications),
(b) upper and lower limits on shipments of a product to a particular plant,
(c) specifications of the product weights for performance measures in the distribution centers,
(d) piecewise linear approximation for nonlinear costs,
(e) the ability to locate plants and distribution centers,
(f) inclusion of capacity constraints of the products in plants,
(g) conversion of raw materials in activities of one or two levels,
(h) additional distribution and production levels.

Reference [7] suggested the inclusion of additional elements in facility location models such as the inclusion of new objectives (maximum return of investment) and decisions regarding the selection of equipment for new installations. In discrete location problems, selecting sites to establish new facilities in is restricted to a finite set of places available for the location. The simplest example of this approach is where $p$ facilities must be selected to minimize the total distance (weighted) or the costs to supply customer demand.

This is a classic problem called the $p$-median problem, which has been extensively studied by [8–11]. This problem assumes that all candidate sites are equivalent in terms of installation cost for a new facility. When this is not the case, the objective function is extended with a term for the fixed cost of location, and as a result, the number of facilities to be open is an endogenous decision. This new approach is known in the literature as the uncapacitated facility location problem (UFLP). There are many references to this problem like in [12–14]. In both cases, the $p$-median and UFLP, each client is assigned to an open facility that minimizes the allocation cost. One of the most important extensions to the UFLP is the capacitated facility location problem (CFLP) in which exogenous values are considered to maximize the demand that can be supplied for each potential site. In this case, the closest assignment property is no longer valid as proposed by [15–17].

The models above have several common characteristics as follows:

(a) single period of planning horizon,
(b) deterministic parameters (demand, costs),
(c) single product,
(d) a single type of facility,
(e) decisions of location-allocation.

Clearly, these models are insufficient to handle a realistic facility location scenario. Therefore, many extensions of this basic problem have been proposed and widely studied [18]. A crucial issue in many practical problems of localization is to consider the existence of different types of facilities [19], each playing a specific role (production or storage), and a natural flow of the material (i.e., a hierarchy) between them [20]. Each set of facilities of the same type is usually denoted by a level or echelon in the hierarchy of facilities.

Part of the planning processes in SCM aim at finding the best possible supply chain configuration so that all operations can be performed in an efficient way. The coming back to the capacitated facility location problem (CFLP) is a well-known combinatorial optimization problem. It consists in deciding which facilities to open from a given potential set, and how to assign customers to those facilities. The objective is minimizing total fixed and shipping cost. Applications of the CFLP include location and distribution planning, lot sizing in production planning, and telecommunication network design as mentioned by [21]. Numerous heuristics and exact algorithms for the CFLP have been proposed in the literature. Heuristic solution methods as well as approximation algorithms were proposed by [22–24]. Tabu Search methods for the related $p$-median problem and the CFLP with single source were developed [25, 26]. Exact solution methods based on the Benders decomposition algorithm are considered [27]. Polyhedral results for the CFLP have been obtained by [28]. Reference [29] uses these results in a branch and cut algorithm for the CFLP.

Moreover, several variants of the CFLP have been investigated. Reference [30] formulated a stochastic integer linear programming model for the CFLP with stochastic demands. A branch and cut approach was applied to find the optimal solution of the problem. Reference [31] formulated the mathematical model of the two-echelon single-source CFLP and considered six Lagrangian relaxation based approaches for the solution. In the recent years, many metaheuristic approaches have been applied to combinatorial optimization problems successfully, such as Simulated Annealing (SA), Genetic Algorithms (GAs), Tabu Search (TS), and Ant

Colony Optimization (ACO). Some recent work in this field includes those presented by [32] in which they use an MAX-MIN ant system approach for the design of a supply chain. Reference [33] presented a memetic algorithm for a multistage supply chain problem. Reference [34] proposes a simulated annealing algorithm for an allocation problem. Reference [35] presents a hybrid approach using an artificial bee algorithm (BA) with mixed integer programming (MIP) applied to a large-scale CFLP; BA is applied for the purpose of solving the location problem, and the MIP is applied for the purpose of finding the optimal mathematical problem.

The biobjective location problems are extensions of classic locations problems. These problems are biobjective median, knapsack, quadratic, covering, unconstrained, location-allocation, hub, hierarchical, competitive, network, and undesirable and semidesirable location problems. Considering capacities in location problems, there are capacitated and uncapacitated problems in the literature. For instance, [36] has considered an uncapacitated facility location problem with two maxisum objectives (net profit and profitability of investment) and modeled it as parametric integer program with fractional and linear objectives. Reference [37] has modeled a supply network as a biobjective uncapacitated facility location problem with minisum and maxisum objectives (cost and coverage). In contrast, [38] developed an extension of the capacitated model to deal with locating maternity facilities with minisum objectives (distance traveled and load imbalance). Reference [39] has used a different bicriteria approach to the single hub location/allocation problem. This approach has two objectives; the first has a minisum form (cost), while the second objective (processing time) has two alternative forms.

The Capacitated Fixed Cost Facility Location Problem with Transportation Choices (CFCLP-TC) proposed by [2] is an extension of the CFLP with a biobjective mixed-integer program approach (cost and time). This approach considers several alternatives to transport the product from one facility to the other in each echelon of the network. At difference from similar works in the literature, the aim here is to provide the decision maker with a set of nondominated alternatives to allow him deciding. Some qualitative information only known by the decision maker may motivate the selection of one of those alternatives.

## 3. Problem Description

The Capacitated Fixed Cost Facility Location Problem with Transportation Choices (CFCLP-TC) proposed by [2] is based on a two-echelon system for the distribution of one product in a single time period. In the first step, the product is sent from manufacturing plants ($i$) to distribution centers ($j$).

The second step corresponds to the flow of product from distribution centers ($j$) to customers ($k$). In this problem, the number and location of plants ($i$) and customers ($k$) are known a priori. This includes a further decision on the selection of channels of transportation between facilities using a bi-objective approach that simultaneously minimizes the time of transportation of product from plants to customers and the combined costs of locating facilities and transportation. This solution approach builds a set of alternative nondominated solutions for the decision maker. This problem has a set of possible locations for the opening of distribution centers ($j$) and their number is not defined. Each candidate site has a fixed cost for opening a facility, and each site has limited capacity. Manufacturing plants have limited capacity and their production is sent from each plant to the distribution centers.

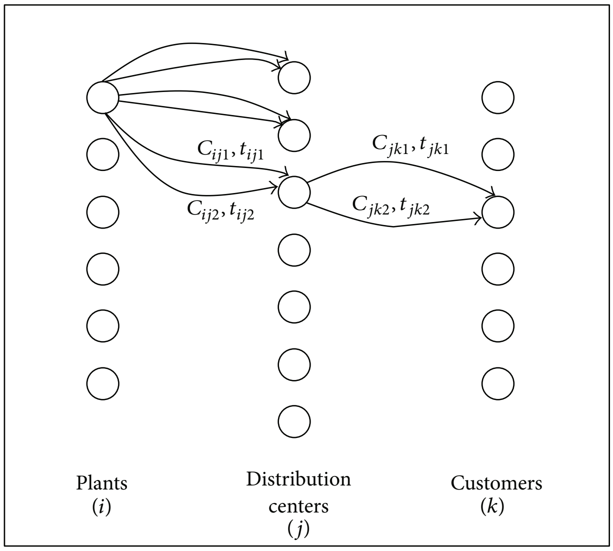



Figure 1: Schematic of the Capacitated Fixed Cost Facility Location Problem with Transportation Choices (CFCLP-TC).

An important feature of the problem is to consider various alternatives for transportation of product from one facility to another in each step of the network. Each option represents a type of service with associated costs and time parameters. The existence of third-party logistics companies (3PL) causes that different transportation services are available in the market. The alternatives are generated by the supply of different companies, the availability of different types of services (urgent or regular), and the use of different modes of transportation (truck, train, plane, ship, or intermodal). Commonly, these differences involve an inverse correlation between time and cost; a faster service is more expensive. The outline of the distribution network is displayed in Figure 1.

*3.1. Approach That Allows Supplying from Multiple Distribution Centers.* In the original model it is established as a restriction that each customer is served by a single source (distribution center ($j$)). At this point, we ask what would happen if the delivery to customers from multiple sources is allowed? The main idea of this variation is to allow customers to be supplied in some cases by more than one source (distribution center ($j$)).

The implementation of this variation to the original model will be evaluated over the objective functions (1) and (2) to determine which has a better Pareto front. Also we study the behavior of the time required to obtain results compared with that required in the original model.

*Sets*

$I$: Set of plants $i$

$J$: Set of potential distribution centers $j$

$K$: Set of customers $k$

$LP_{ij}$: Set of arcs $l$ between nodes $i$ and $j$; $i \in I, j \in J$

$LW_{jk}$: Set of arcs $l$ between nodes $j$ and $k$: $j \in J, k \in K$.

*Parameters*

$CP_{ijl}$: Cost of transporting one unit of product from plant $i$ to distribution center $j$ using the arc $l$; $i \in I, j \in J, l \in LP_{ij}$

$CW_{jkl}$: Cost of transporting one unit of product from distribution center $j$ to customer $k$ using the arc $l$; $j \in J, k \in K, l \in LW_{jk}$

$TP_{ijl}$: Time for transporting any quantity of product from plant $i$ to the distribution center $j$ using arc $l$; $i \in I$, $j \in J$, $l \in LP_{ij}$

$TW_{jkl}$: Time for transporting any quantity of product from distribution center $j$ to customer $k$ using arc $l$; $j \in J$, $k \in K$, $l \in LW_{jk}$

$MP_i$: Capacity of plant $i$; $i \in I$

$MW_j$: Capacity of distribution center $j$; $j \in J$

$D_k$: Demand of customer $k$; $k \in K$

$F_j$: Fixed cost for opening distribution center $j$; $j \in J$.

*Decision Variables*

$X_{ijl}$: Quantity transported from plan $i$ to distribution center $j$ using arc $l$; $i \in I$, $j \in J$, $l \in LP_{ij}$

$Y_{jkl}$: Quantity transported from distribution center $j$ to customer $k$ using arc $l$; $j \in J, k \in K, l \in LW_{jk}$

$Z_j$: Binary variable equal to 1 if distribution center $j$ is open and equal to 0 otherwise; $j \in J$

$A_{ijl}$: Binary variable equal to 1 if arc $l$ is used to transport product from plant $i$ to distribution center $j$ and equal to 0 otherwise; $i \in I$, $j \in J, l \in LP_{ij}$

$B_{jkl}$: Binary variable equal to 1 if arc $l$ is used to transport product from distribution center $j$ to customer $k$ and equal to 0 otherwise; $j \in J$, $k \in K$, $l \in LW_{jk}$.

*Auxiliary Variables*

$T$: A variable that computes the longest time that takes sending product from any plant to any customer

$E_j^1$: Longest time in the first echelon of the supply chain for active distribution center $j$; that is, $E_j^1 = \max_{i,l}(TP_{ijl}A_{ijl})$; $i \in I$, $j \in J, l \in LP_{ij}$

$E_j^2$: Longest time in the second echelon of the supply chain for active distribution center $j$; that is, $E_j^2 = \max_{k,l}(TW_{jkl}B_{jkl})$; $j \in J, k \in K, l \in LW_{jk}$.

*Model 1.* Consider $\min(f_1, f_2)$

$$f_1 = \sum_{i\in I}\sum_{j\in J}\sum_{l\in LP_{ij}} CP_{ijl}X_{ijl} + \sum_{j\in J}\sum_{k\in K}\sum_{l\in LW_{jk}} CW_{jkl}Y_{jkl} + \sum_{j\in J} F_j Z_j, \tag{1}$$

$$f_2 = T, \tag{2}$$

$$T - E_j^1 - E_j^2 \geq 0 \quad \forall j \in J, \tag{3}$$

$$E_j^1 - TP_{ijl}A_{ijl} \geq 0 \quad \forall i \in I, j \in J, l \in LP_{ij}, \tag{4}$$

$$E_j^2 - TW_{jkl}B_{jkl} \geq 0 \quad \forall j \in J, k \in K, l \in LW_{jk}, \tag{5}$$

$$\sum_{j\in J}\sum_{l\in LW_{jk}} Y_{jkl} = D_k \quad \forall k \in K, \tag{6}$$

$$\sum_{j\in J}\sum_{l\in LP_{ij}} X_{ijl} = MP_i \quad \forall i \in I, \tag{7}$$

$$MW_j Z_j - \sum_{j\in J}\sum_{l\in LW_{jk}} Y_{jkl} \geq 0 \quad \forall j \in J, \tag{8}$$

$$\sum_{j\in J}\sum_{l\in LP_{ij}} X_{ijl} - \sum_{j\in J}\sum_{l\in LW_{jk}} Y_{jkl} = 0 \quad \forall j \in J, \tag{9}$$

$$\sum_{l\in LP_{ij}} A_{ijl} \leq 1 \quad \forall i \in I, j \in J, \tag{10}$$

$$\sum_{l\in LW_{jk}} B_{jkl} \leq 1 \quad \forall j \in J, k \in K, \tag{11}$$

$$X_{ijl} - A_{ijl} \geq 0 \quad \forall i \in I, j \in J, l \in LP_{ij}, \tag{12}$$

$$Y_{jkl} - B_{jkl} \geq 0 \quad \forall j \in J, k \in K, l \in LW_{jk}, \tag{13}$$

$$MP_i A_{ijl} - X_{ijl} \geq 0 \quad \forall i \in I, j \in J, l \in LP_{ij}, \tag{14}$$

$$MW_j B_{jkl} - Y_{jkl} \geq 0 \quad \forall j \in J, k \in K, l \in LW_{jk}, \tag{15}$$

$$\sum_{i\in I}\sum_{l\in LP_{ij}} A_{ijl} - Z_j \geq 0 \quad \forall j \in J, \tag{16}$$

$$T, E_j^1, E_j^2, X_{ijl}, Y_{jkl} \geq 0 \quad \forall i \in I, \quad j \in J, \quad k \in K, \quad l \in LP_{ij}, \quad l \in LW_{jk} \tag{17}$$

$$Z_j, A_{ijl}, B_{jkl} \in \{0, 1\} \quad \forall i \in I, \quad j \in J, \quad k \in K, \quad l \in LP_{ij}, \quad l \in LW_{jk}. \tag{18}$$

In this model, the objective function (1) minimizes the transportation costs and the cost of opening the distribution

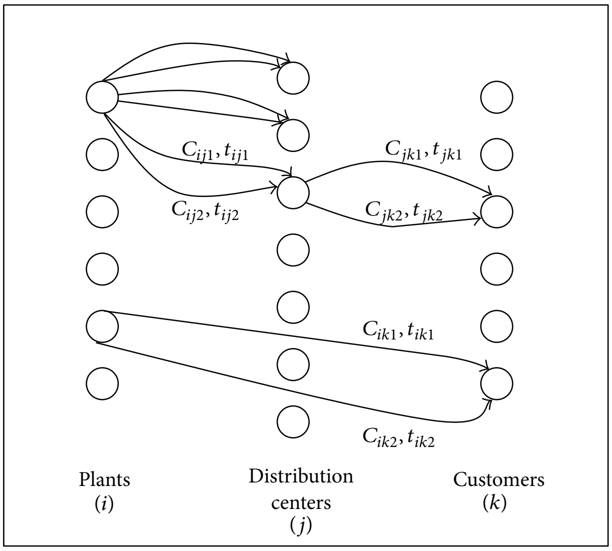



Figure 2: Schematic of the Capacitated Fixed Cost Facility Location Problem with Transportation Choices (CFCLP-TC) with variations in the direct flow between plants ($i$) and customers ($k$).

centers. The objective function (2) minimizes the time of transportation. Restriction (3) calculates the longest total time between ($i$) and ($k$). The restrictions (4) and (5) allow calculating the transportation time between ($i$) and ($j$) and then from ($j$) to ($k$). The restriction (6) allows the satisfaction of the demands of each client. The restriction (7) impedes to exceed the capacity of each plant ($i$). The restriction (8) impedes to exceed the capacity of the distribution centers ($j$). The restriction (9) allows the balance of flow between ($i$)-($j$) and ($j$)-($k$). Restrictions (10) and (11) provide that the transportation of the material can only be done through a single arc. Restrictions (12) and (13) provide that an arc will be inactive if there is no flow through it. Equations (14) and (15) state the shipment of product only through active arcs. Restriction (16) provides that the distribution centers ($j$) having no product flow through them must be closed. Restrictions (17) and (18) provide a definition of the variables in the model.

*3.2. Approach Allowing the Direct Flow between Plants and Customers.* This variation suggests that some plants ($i$) supply customers in a direct way, that is, without necessarily passing through the distribution centers ($j$). The main reason is because in some cases customers ($k$) may be closer to the plants ($i$) than the distribution centers ($j$). This could make the flow more efficient in terms of cost and time without considering distribution centers ($j$). As in the previous variation, we want to evaluate the efficiency of this approach through the construction of a new Pareto front and the performance evaluation of the computational time required. Figure 2 shows the outline of this proposal.

The mathematical model for the variation that allows the direct flow between plants ($i$) and costumers ($k$) in some cases is as follows.

*Sets*

$LV_{ik}$: Set of arcs $l$ between nodes $i$ and $k$: $i \in I$, $k \in K$

*Parameters*

$CV_{ikl}$: Cost of sending one unit of product from plant $i$ to customer $k$ using the arc $ikl$: $i \in I$, $k \in K$, $l \in LV_{ik}$

$TV_{ikl}$: Time to transport any quantity of products of plant $i$ to customer $k$ using the arc $ikl$; $i \in I$, $k \in K$, $l \in LV_{ik}$.

*Decision Variables*

$V_{ikl}$: Quantity transported from plant $i$ to customer $k$ using the arc $ikl$; $i \in I, k \in K, l \in LV_{ik}$

$G_{ikl}$: Binary variable equal to 1 if the arc $ikl$ is used to transport product from plant $i$ to customer $k$ and equals 0 otherwise $i$ and $k$; $i \in I$, $k \in K$, $l \in LV_{ik}$.

*Auxiliary Variables*

$E3$: The maximum time it takes to ship the $i$ to $k$, $E_3 = \max_{k,l}(TV_{ikl}G_{ikl}); i \in I, k \in K, l \in LV_{ik}$.

*Model 2.* Consider $\min(f_1, f_2)$

$$f_1 = \sum_{i\in I}\sum_{j\in J}\sum_{l\in LP_{ij}} CP_{ijl}X_{ijl} + \sum_{j\in J}\sum_{k\in K}\sum_{l\in LW_{jk}} CW_{jkl}Y_{jkl} + \sum_{i\in I}\sum_{k\in K}\sum_{l\in LV_{ik}} CV_{ikl}V_{ikl} + \sum_{j\in J} F_j Z_j \quad (19)$$

(2), (3), (4), (5), (8), (9), (11),

(12), (13), (14), (15), (16), (17),

$$E3 - TV_{ikl}G_{ikl} \geq 0 \quad \forall i \in I,\ k \in k,\ l \in LV_{ik} \quad (20)$$

$$T - E3 \geq 0 \quad \forall j \in J, \quad (21)$$

$$\sum_{j\in J}\sum_{l\in LW_{jk}} Y_{jkl} + \sum_{i\in I}\sum_{l\in LV_{ik}} V_{ikl} = D_k \quad \forall i \in I,\ k \in K, \quad (22)$$

$$\sum_{j\in J}\sum_{l\in LP_{ij}} X_{ijl} + \sum_{i\in I}\sum_{l\in LV_{ik}} V_{ikl} = MP_i \quad \forall i \in I, k \in K, \quad (23)$$

$$\sum_{i\in I}\sum_{l\in LV_{ik}} G_{ikl} + \sum_{j\in J}\sum_{l\in LW_{jk}} B_{jkl} = 1 \quad \forall i \in I,\ k \in K, \quad (24)$$

$$\sum_{l\in LV_{ik}} G_{ikl} \leq 1 \quad \forall i \in I,\ k \in K, \quad (25)$$

$$V_{ikl} - G_{ikl} \geq 0 \quad \forall i \in I,\ k \in K,\ l \in LV_{ik}, \quad (26)$$

$$MP_i G_{ikl} - V_{ikl} \geq 0 \quad \forall i \in I, k \in K, l \in LV_{ik}, \quad (27)$$

$$T, E_j^1, E_j^2, E3, X_{ijl}, Y_{jkl}, G_{ikl} \geq 0 \quad \forall i \in I, \quad j \in J, \quad k \in K, \quad l \in LP_{ij}, \quad l \in LW_{jk}, \quad l \in LV_{ik}, \quad (28)$$

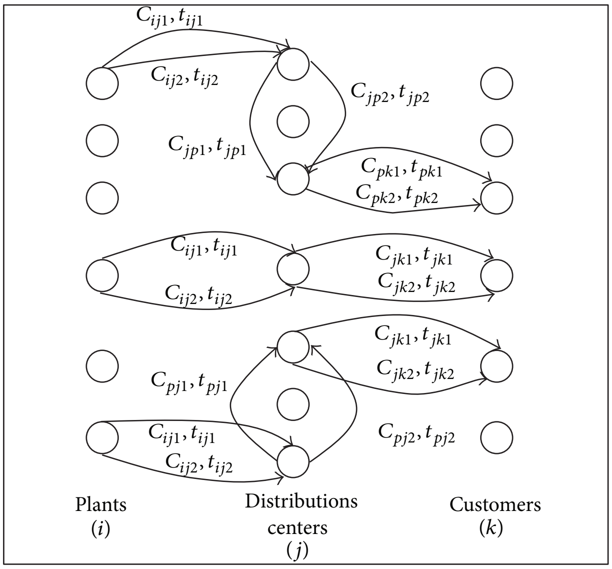



FIGURE 3: Schematic of the Capacitated Fixed Cost Facility Location Problem with Transportation Choices (CFCLP-TC) flow between distribution centers.

$$Z_j, A_{ijl}, B_{jkl}, V_{ikl} \in \{0, 1\}$$
$$\forall i \in I, \quad j \in J, \quad k \in K, \quad l \in LP_{ij}, \tag{29}$$
$$l \in LW_{jk}, \quad l \in LV_{ik}.$$

In this new formulation (19) replaces (1) in the original formulation as the objective function that searches for the best cost. Equation (20) is added to the original model and this calculates the time between the plants ($i$) and customers ($k$). Equation (21) assigns the longest time between plants ($i$) and customers ($k$) to variable $T$. Equation (22) replaces (6) in the original model and is aimed at satisfying the customer demand ($k$). Equation (23) replaces (7) in the original model and this requires that the transported amount from ($i$) to ($k$) does not exceed the capacity of the plant ($i$). Equation (24) is added to the model, and this ensures that customers ($k$) can be supplied only by a single source. Equation (25) is added to the original model and states that the transportation of the material can only be done through a single arc. Equation (26) is added to the original model to establish that an arc is inactive whether its flow is zero. Equation (27) is added to the original model and states that the product flow will be made only through active arcs. Equations (28) and (29) replace (18) and (19) in the original model and set the domain of the variables.

*3.3. Approach That Allows Flow between Distribution Centers.* This variation suggests the possibility of exchange of goods between distribution centers ($j$). For some alternatives this variation would compete directly with the flow of plants ($i$), distribution centers ($j$), plants ($k$) with the alternative flow of plants ($i$), distribution centers ($j$), distribution centers ($p$), and plants ($k$). The main idea is that for some cases it is cheaper and faster to send a product from a distribution center ($j$) to other distribution centers ($p$) and then send it to customers ($k$). Figure 3 shows the outline of this proposal.

The mathematical model allowing flow between distribution centers, ($j$)-($p$), is as follows.

*Sets*

$LS_{jp}$: Set of arcs $l$ between nodes $j$-$p$: $j \in J$

$LS_{pj}$: Set of arcs $l$ between nodes $p$-$j$: $j \in J$.

*Parameters*

$CS_{jpl}$: Cost of sending one unit of product between distribution centers $j$ using arc $jpl$: $j \in J$, $l \in LS_{jp}$

$CS_{pjl}$: Cost of sending one unit of product between distribution centers $p$ using arc $pjl$: $j \in J$, $l \in LS_{pj}$

$TS_{jpl}$: Time to transport any quantity of products between distribution centers $j$ using the arc $jpl$; $j \in j$, $l \in LS_{jp}$

$TS_{pjl}$: Time to transport any quantity of products between distribution centers $p$ using the arc $pjl$; $j \in j$, $l \in LS_{pj}$.

*Decision Variables*

$R_{jpl}$: Quantity transported from $j$ to$p$ using arc $jpl$; $j \in J$, $l \in LS_{jp}$

$R_{pjl}$: Quantity transported from $p$ to $j$ using arc $pjl$; $j \in J$, $l \in LS_{pj}$

$C_{jpl}$: Binary variable equal to 1 if the arc $jpl$ is used to transport product between distribution centers $(j-p)$ and equals 0 otherwise; $j \in J$, $p \in P$, $l \in LS_{jp}$

$C_{pjl}$: Binary variable equal to 1 if the arc $pjl$ is used to transport product between distribution centers $(p-j)$ and equals 0 otherwise; $j \in J$, $p \in P$, $l \in LS_{pj}$.

*Auxiliary Variables*

$T$: Maximum time it takes to ship the product using the alternative $i$, $j$, $p$, $k$, $E_3 = \max(TP_{ijl}A_{ijl}) + \max(TS_{jpl}C_{jpl}) + \max(TW_{jkl}B_{jkl})$; $i \in I, j \in J, l \in LP_{ij}, l \in LS_{jp}, k \in K, l \in LW_{jk}$

$E_3$: Alternatively time $i$-$j$-$p$-$k$, $E_3 = \underset{ijpk}{\mathrm{Max}}\,[TP_{ijl} + TS_{pjl} + TW_{jkl}]$

$M$: A very large positive value

$\delta_{jp}$: Binary variable equal to 1 if the arc $jpl$ is used to transport product from $j$ to $p$ and equals 0 otherwise$j \in J, l \in LS_{jp}$

$\delta_{pj}$: Binary variable equal to 1 if the arc $pjl$ is used to transport product from $p$ to $j$ and equals 0 otherwise $\in J, l \in LS_{pj}$.

*Model 3.* Consider $\min(f_1, f_2)$

$$f_1 = \sum_{i\in I}\sum_{j\in J}\sum_{l\in LP_{ij}} CP_{ijl}X_{ijl} + \sum_{j\in J}\sum_{k\in K}\sum_{l\in LW_{jk}} CW_{jkl}Y_{jkl} + \sum_{j\in J}\sum_{p\in P}\sum_{l\in LS_{jp}} C_{jpl}R_{jpl} + \sum_{j\in J}\sum_{p\in P}\sum_{l\in LS_{pj}} C_{pjl}R_{pjl} + \sum_{j\in J} F_j Z_j \quad (30)$$

(2), (3), (4), (5), (6), (7), (8), (10),

(11), (12), (13), (14), (15), (16),

$$T - E_3 \geq 0 \quad \forall j \in J, \quad (31)$$

$$\sum_{l\in LS_{jp}} C_{jpl} \leq M\left(1-\delta_{jp}\right) \quad \forall j \in J, p \in P, \quad (32)$$

$$\sum_{l\in LS_{pj}} C_{pjl} \leq M\left(1-\delta_{pj}\right) \quad \forall j \in J, p \in P, \quad (33)$$

$$\sum_{j\in J}\sum_{l\in LS_{jp}} C_{jpl} + \sum_{p\in P}\sum_{l\in LS_{jp}} C_{pjl} \leq 1 \quad \forall j \in J, p \in P, \quad (34)$$

$$E_3 - \sum_{l\in LP_{ij}} A_{ijl}TP_{ijl} + \sum_{l\in LS_{jp}} C_{jpl}TS_{jpl} + \sum_{l\in LW_{pk}} B_{pkl}TW_{pkl} \geq -M\delta_{jp} \quad (35)$$

$$\forall i \in I, \quad j \in J, \quad p \in P, \quad k \in K,$$

$$E_3 - \sum_{l\in LP_{ij}} A_{ipl}TP_{ipl} + \sum_{l\in LS_{pj}} C_{pjl}TS_{pjl} + \sum_{l\in LW_{pk}} B_{jkl}TW_{jkl} \geq -M\delta_{pj}$$

$$\forall i \in I, \quad j \in J, \quad p \in P, \quad k \in K, \quad (36)$$

$$MW_j Z_j - \left(\sum_{j\in J}\sum_{l\in LW_{jk}} Y_{jkl} + \sum_{j\in LS_{jp}} R_{jpl} - \sum_{j\in LS_{pj}} R_{pjl}\right) \geq 0 \quad \forall j \in J, \quad (37)$$

$$\sum_{i\in I}\sum_{l\in LP_{ij}} X_{ijl} + \sum_{p\in P}\sum_{l\in LS_{jp}} R_{pjl} - \sum_{k\in K}\sum_{l\in LW_{jk}} Y_{jkl} - \sum_{p\in P}\sum_{l\in LS_{jp}} R_{jpl} = 0 \quad \forall j \in J, \quad (38)$$

$$\sum_{l\in LS_{jp}} C_{jpl} \leq 1 \quad \forall j \in J, \quad (39)$$

$$\sum_{l\in LS_{pj}} C_{pjl} \leq 1 \quad \forall j \in J, \quad (40)$$

$$\sum_{p\in P}\sum_{l\in LS_{jp}} C_{jpl} + \sum_{p\in P}\sum_{l\in LS_{jp}} C_{pjl} \leq 1 \quad \forall j \in J, \quad (41)$$

$$R_{jpl} - C_{jpl} \geq 0 \quad \forall j \in J,\ l \in LS_{jp}, \quad (42)$$

$$R_{pjl} - C_{pjl} \geq 0 \quad \forall j \in J,\ l \in LS_{pj}, \quad (43)$$

$$MW_j C_{jpl} - R_{jpl} \geq 0 \quad \forall j \in J,\ l \in LS_{jp}, \quad (44)$$

$$MW_j C_{pjl} - R_{pjl} \geq 0 \quad \forall j \in J,\ l \in LS_{pj} \quad (45)$$

$$T, E_j^1, E_j^2, E_3, X_{ijl}, Y_{jkl}, R_{jpl} \geq 0 \quad \forall i \in I, \quad j \in J, \quad k \in K, \quad l \in LP_{ij}, \quad l \in LW_{jk}, \quad l \in LS_{jp}, \quad (46)$$

$$Z_j, A_{ijl}, B_{jkl}, C_{jpl}, \delta_{jp} \in \{0,1\} \quad \forall i \in I, \quad j \in J, \quad k \in K, \quad l \in LP_{ij}, \quad l \in LW_{jk}, \quad l \in LS_{jp}. \quad (47)$$

In this new formulation (30) replaces (1) in the original model as an objective function that looks for the best possible cost. Equation (31) is added to the original model and calculates the longest time considering the flow $(j)$-$(p)$. Equations (32) and (33) are added to the original set as a condition if-then to determine the longest time in the flow of product from $(i)$ to $(k)$, considering the flow between $(j)$ and $(p)$. Equation (34) is added to the original model and this limits multiple routing between distribution centers $(j)$. Equations (35) and (36) are added to the original model and complete the if-then condition to calculate the longest time in the flow of product from $(i)$ to $(k)$ considering the flow of $(j)$-$(p)$. Equation (37) replaces (8) in the original model and states that the amount transported from $(j)$ to $(k)$ does not exceed the capacity of distribution center $(j)$ by considering the flow between distribution centers $(j)$-$(p)$. Equation (38) replaces (9) in the original model and provides the flow balance in the distribution centers. Equations (39) and (40) are added to the original model and these provide that the transportation of the material can only be done through a single arc. Equation (41) is added to the original model and states that the exchange of product between $(j)$ and $(p)$ does not create a cycle. Equations (42) and (43) provide that an arc will be inactive if there is no flow through it. Equations (44) and (45) are added to the original model and establish that the distribution centers $(j)$, having no product flow through them, will be closed. Equations (46) and (47) replace (18) and (19) in the original model and set the domain of the variables in the model.

## 4. Computational Experiment

For the process of computational experiments, five sets of instances of each size were used as shown in Table 1.

Table 1: Instances sizes.

| Instances sizes | Integer variables in the original model | Integer variables in the approach without single source | Integer variables in the approach allowing direct flow between ($i$) and ($k$) | Integer variables in the approach allowing flow between distribution centers ($j$)-($p$) |
|---|---|---|---|---|
| 5_5_5_2 | 105 | 105 | 205 | 355 |
| 5_5_5_5 | 255 | 255 | 505 | 805 |
| 5_10_10_2 | 320 | 320 | 510 | 1310 |
| 5_10_15_2 | 410 | 410 | 710 | 1410 |
| 5_10_20_2 | 510 | 510 | 910 | 1510 |

The encoding of the instance sizes is as follows. The first index indicates the number of plants ($i$), the second index indicates the number of distribution centers ($j$), the third index indicates the number of customers ($k$), and finally the fourth index indicates the number of arcs between nodes in each echelon. In each size, 5 instances were tested.

*4.1. Instance Design.* The instances used in the experimental process were initially proposed by [2, 3]. These instances were used to test the original model and the variation that allows multiple sources of supply for customers.

For the variation that permits direct supply from the plants ($i$) to customers ($k$), these instances were extended to accommodate this approach. The times for the arcs of the alternative ($i$)-($k$) were generated at random on the basis of a normal distribution, with values ranging from 5 to 50. The aim of these values is that they are competitive with those of the original instance so that the model can choose to select sometimes the direct flow of ($i$)-($k$) and in other cases to select the flow ($i$)-($j$)-($k$). We consider that time and cost for the alternative ($i$)-($k$) are negatively correlated. The unit cost of transportation in the flow ($i$)-($k$) is a floating point variable calculated as follows:

$$\mathrm{Cost}_1 = \frac{(7) * (50)}{\mathrm{Time}}. \tag{48}$$

This factor of "7" was determined after some experiments with some instances in such a way that the different solutions along the Pareto front included direct paths ($i$)-($k$) and indirect paths ($i$)-($j$)-($k$). This was done to impede that the model would prefer a certain type of path.

To test the model that allows the exchange of product between distribution centers, we intended to use the original data extended by adding new times and costs for the arcs that allow the exchange between distribution centers ($j$). With these data the model did not select arcs ($j$)-($p$), because in fact the solutions had few open distribution centers ($j$). To promote the opening of more distribution centers the fixed opening cost should be reduced in the new instances. With this change, we expected that some arcs ($j$)-($p$) could be used in the solutions. Therefore, the extension to the original instance is as follows: the fixed cost of distribution center ($j$) was generated as a random variable with integer values between 10,000 and 15,000.

These values will force the model to have more open distribution centers so that the product flow between them is more likely. From this point, the instances were expanded, creating additional time for flows ($j$)-($p$) at random with a range of 1 to 5. The costs of the new alternatives are generated using the following relationship:

$$\mathrm{Cost}_2 = \left(\frac{50}{\mathrm{Time}}\right) * \left(\frac{1}{10}\right). \tag{49}$$

This factor of "1/10" was determined after some experiments with some instances in such a way that the different solutions along the Pareto front included ($i$)-($j$)-($k$) paths and ($i$)-($j$)-($p$)-($k$) paths. This was done to impede that the model would prefer a certain type of path.

To perform the computational experiment we used a computer equipped with the following features: Workstation with Intel (R) Xenon (R) CPU X5550 2.67 GHz with 12 GB of RAM and 64-bit operating system (Windows 7). The implementation of the models was performed in GAMS 12 and solved with CPLEX 23.6.2.

*4.2. Metrics.* To make the comparison of the new Pareto fronts with the original ones, the metric $R_{\mathrm{pos}}(P_i)$ proposed by [4] was used. Additionally, we registered the average number of Pareto-optimal solutions in each front. To calculate the $R_{\mathrm{pos}}(P_i)$, consider that $P_1$ and $P_2$ are the sets of Pareto-optimal solutions obtained from each model and $P$ is the union of the sets of Pareto-optimal solutions (i.e., $P = P_1 \cup P_2$) such that it includes only nondominated solutions $Y$'s. The ratio of Pareto-optimal solutions in $P_i$ that are not dominated by any other solutions in $P$ is calculated as follows:

$$R_{\mathrm{pos}}(P_i) = \frac{\left|P_i - \{X \in P_i \mid \exists T \in P : Y \prec X\}\right|}{P_i}, \tag{50}$$

where $Y \prec X$ means that the solution $X$ is dominated by solution $Y$. The higher the ratio $R_{\mathrm{pos}}(P_i)$ is, the better the solution set $P_i$ is. Similarly, we used the metrics proposed by [3] called $D_{\mathrm{avg}}$ and $D_{\mathrm{min}}$. These were developed to give practical meaning to the comparison of sets point by point. The discretization of objective $f_2$ and the number of objectives allow proceeding as follows for a pair of sets $S_1$ and $S_2$.

Let $f_1$ and $f_2$ be the objective functions of the problem and $S_1$ and $S_2$ the sets of nondominated solutions to be compared. By discretization of function $f_2$ we can construct

Table 2: Results of the first variation.

| Size | With single source | | | Without single source | | | $D_{min}$ | Time with single source (seconds) | Time without single source (seconds) |
|---|---|---|---|---|---|---|---|---|---|
| | Instance | $\|S_i\|$ | $R_{pos}(S_i)$ | $\|S_i\|$ | $R_{pos}(S_i)$ | $D_{avg}$ | | | |
| 5_5_5_2 | 1 | 23 | 0.71875 | 32 | 1 | 0.99924159 | 0.98804728 | 8.838 | 9.068 |
| | 2 | 13 | 0.46428571 | 28 | 1 | 0.99736445 | 0.98254824 | 9.197 | 11.052 |
| | 3 | 15 | 0.53571429 | 28 | 1 | 0.99803935 | 0.99272805 | 9.606 | 10.751 |
| | 4 | 18 | 0.5625 | 32 | 1 | 0.99726561 | 0.97547912 | 14.514 | 15.811 |
| | 5 | 24 | 0.96 | 25 | 1 | 0.99932249 | 0.98894956 | 9.006 | 9.565 |
| 5_5_5_5 | 1 | 34 | 0.89473684 | 38 | 1 | 0.99955038 | 0.99126315 | 40.998 | 59.376 |
| | 2 | 8 | 0.2 | 40 | 1 | 0.99470283 | 0.97594721 | 55.89 | 68.319 |
| | 3 | 25 | 0.64102564 | 39 | 1 | 0.99933261 | 0.98922985 | 68.772 | 87.766 |
| | 4 | 30 | 0.76923077 | 39 | 1 | 0.99938233 | 0.99327893 | 55.651 | 93.328 |
| | 5 | 37 | 0.94871795 | 39 | 1 | 0.99993551 | 0.99766135 | 41.012 | 52.617 |
| 5_10_10_2 | 1 | 0 | 0 | 37 | 1 | 0.96917717 | 0.92948071 | 356.355 | 2294. 165 |
| | 2 | 0 | 0 | 39 | 1 | 0.98128174 | 0.96600743 | 193.005 | 371.465 |
| | 3 | 0 | 0 | 37 | 1 | 0.98124332 | 0.96250708 | 417.547 | 2212.05 |
| | 4 | 2 | 0.05 | 40 | 1 | 0.97790924 | 0.94280589 | 233.841 | 813.129 |
| | 5 | 0 | 0 | 37 | 1 | 0.96969167 | 0.91712171 | 259.442 | 1674.431 |
| 5_10_15_2 | 1 | 0 | 0 | 36 | 1 | 0.97493663 | 0.93690112 | 8042.069 | 29027.322 |
| | 2 | 0 | 0 | 36 | 1 | 0.98081129 | 0.95607975 | 11207.506 | 2668.63 |
| | 3 | 0 | 0 | 35 | 1 | 0.98012085 | 0.96681874 | 16262.752 | 8218.611 |
| | 4 | 0 | 0 | 37 | 1 | 0.98183982 | 0.96352299 | 23225.921 | 10991.189 |
| | 5 | 0 | 0 | 39 | 1 | 0.98417036 | 0.96930275 | 8734.642 | 6810.192 |
| 5_10_20_2 | 1 | 0 | 0 | 35 | 1 | 0.98737854 | 0.97111859 | 149028.323 | 364.128 |
| | 2 | 0 | 0 | 37 | 1 | 0.98613018 | 0.97454929 | 112315.372 | 24367.128 |
| | 3 | 0 | 0 | 35 | 1 | 0.99016256 | 0.9824318 | 246814.212 | 38991.51 |
| | 4 | 0 | 0 | 37 | 1 | 0.98522424 | 0.973581 | 240520 | 39919.572 |
| | 5 | 0 | 0 | 37 | 1 | 0.98692727 | 0.97677506 | 78406 | 32163.298 |

a set $T$, such that its elements are those values of $f_2$ that exist in $S_1$ and $S_2$,

$$T = \{f_2(s) \vee f_2(s'), s \in S_1, s' \in S_2 \mid \exists f_1(s) \wedge \exists f_1(s') \wedge f_2(s) = f_2(s')\}. \tag{51}$$

Then $D_{avg}$ computes an average rate deviation of the objective function $f_1$ for each value of $f_2$, which is in the set $T$,

$$D_{avg} = \frac{\sum_{t\in T}\left(\left(f_1(s) : f_2(s) = t\right) / \left(f_1(s') : f_2(s') = t\right)\right)}{|T|} \quad \forall s \in S_1, \quad s' \in S_2,$$

$$D_{min} = \min_{t\in T} \frac{\left(f_1(s) : f_2(s) = t\right)}{\left(f_1(s') : f_2(s') = t\right)} \quad \forall s \in S_1, \quad s' \in S_2. \tag{52}$$

The metric $D_{avg}$ indicates the quality of a set compared to another. The following relationship can be established:

$$\text{If } D_{avg} \begin{cases} < 1 & S_1 \text{ is better than } S_2 \\ > 1 & S_1 \text{ is worse than } S_2 \\ = 1 & S_1 \text{ is similar to } S_2. \end{cases} \tag{53}$$

It is important to establish that an important parameter is the computational time required to solve each model for an instance.

*4.3. Results.* Table 2 shows the comparison results for the solution of the original model and the first variation (without single source constraint) of the CFCLP-TC problem. In this evaluation five instances are presented with metrics $R_{pos}$, $D_{avg}$, and $D_{min}$ and the processing time in seconds of each one.

Table 2 shows that in all cases $D_{avg}$ is less than 1; this indicates a superior quality of the Pareto fronts of the variation compared to the original model. $D_{min}$ in all cases indicates a small difference between both fronts, since the values are close to 1. The values of $R_{pos}$ in the original model are on average 64% for instance 5-5-5-2, 69% for instances

Table 3: Results of the second variation.

| Size | Instance | Original model $\lvert S_i \rvert$ | Original model $R_{pos}$ $(S_i)$ | Direct flow between $(i)$ and $(k)$ $\lvert S_i \rvert$ | Direct flow between $(i)$ and $(k)$ $R_{pos}$ $(S_i)$ | $D_{avg}$ | $D_{min}$ | Time with original model (seconds) | Time direct flow between $(i)$ and $(k)$ (seconds) |
|---|---|---|---|---|---|---|---|---|---|
| 5_5_5_2 | 1 | 27 | 0.87096774 | 31 | 1 | 0.9975774 | 0.95965946 | 8.838 | 9.263 |
| | 2 | 15 | 0.42857143 | 35 | 1 | 0.98096094 | 0.91788812 | 9.197 | 10.276 |
| | 3 | 18 | 0.54545455 | 33 | 1 | 0.96701778 | 0.85386306 | 9.606 | 10.456 |
| | 4 | 1 | 0.03225806 | 31 | 1 | 0.95896192 | 0.91764374 | 14.514 | 12.989 |
| | 5 | 23 | 0.79310345 | 29 | 1 | 0.99362365 | 0.96388563 | 9.006 | 9.018 |
| 5_5_5_5 | 1 | 38 | 0.92682927 | 41 | 1 | 1 | 1 | 40.977 | 72.21 |
| | 2 | 33 | 0.76744186 | 43 | 1 | 0.99936964 | 0.99153492 | 55.504 | 99.961 |
| | 3 | 16 | 0.38095238 | 42 | 1 | 0.98025659 | 0.93238372 | 68.331 | 106.332 |
| | 4 | 35 | 0.81395349 | 43 | 1 | 0.99788515 | 0.97854988 | 55.618 | 105.304 |
| | 5 | 34 | 0.79069767 | 43 | 1 | 0.99839478 | 0.96884197 | 40.962 | 69.114 |
| 5_10_10_2 | 1 | 18 | 0.46153846 | 39 | 1 | 0.98846976 | 0.94669285 | 351.812 | 528.167 |
| | 2 | 32 | 0.7804878 | 41 | 1 | 0.99766759 | 0.97732813 | 192.369 | 293.994 |
| | 3 | 27 | 0.675 | 40 | 1 | 0.99658614 | 0.96356955 | 416.336 | 540.134 |
| | 4 | 31 | 0.775 | 40 | 1 | 0.99855269 | 0.98577745 | 233.123 | 393.202 |
| | 5 | 19 | 0.52777778 | 36 | 1 | 0.98881179 | 0.89438732 | 259.084 | 341.4 |
| 5_10_15_2 | 1 | 2 | 0.555556 | 36 | 1 | 0.91196956 | 0.7597195 | 6803.334 | 444.791 |
| | 2 | 11 | 0.28205128 | 36 | 1 | 0.98962601 | 0.94587785 | 9754.167 | 10448.886 |
| | 3 | 11 | 0.275 | 39 | 1 | 0.97893561 | 0.86283473 | 9105.095 | 6400.084 |
| | 4 | 14 | 0.35 | 40 | 1 | 0.97651925 | 0.87874916 | 12192.357 | 10856.243 |
| | 5 | 18 | 0.15 | 40 | 1 | 0.98074666 | 0.91927882 | 7987.301 | 7619.696 |
| 5_10_20_2 | 1 | 3 | 0.07692308 | 39 | 1 | 0.94659341 | 0.80255363 | 149028.323 | 137238.826 |
| | 2 | 9 | 0.24324324 | 37 | 1 | 0.95456087 | 0.83359301 | 112315.372 | 112106.767 |
| | 3 | 10 | 0.26315789 | 38 | 1 | 0.96070312 | 0.83766502 | 246814.212 | 129141.434 |
| | 4 | 8 | 0.20512821 | 39 | 1 | 0.96055131 | 0.83801063 | 240520.165 | 246519.6 |
| | 5 | 9 | 0.23076923 | 39 | 1 | 0.96889729 | 0.56548166 | 78406.285 | 467645.918 |

5-5-5-5, and 0% for instances 5-10-10-2, 5-10-15-2, and 5-10-20-2. This indicates that in all cases the variation that allows the supply of product to customers without single source constraint has always better Pareto fronts compared with those obtained in the original model.

The processing time for the variation increased on average by 9% for instances 5-5-5-2, increased on average 627% for instances 5-5-5-5, and increased by 76% on average for instances 5-10-10-2 compared to the original model. However for instances 5-10-15-2 time decreased by 16% and by 500% for instances 5-10-20-2.

In Table 3 the results of the comparison between the original model and the variation that allows the direct flow between the plants $(i)$ and customers $(k)$ are shown. It should be noted that, in the majority of the cases, $D_{avg}$ is less than 1; this indicates that the Pareto fronts of the variation are better compared to the original model. $D_{min}$ in all cases indicates the smallest difference comparing both fronts and provides a measure of the difference of the fronts compared.

In relation to $R_{pos}$ it is observed that the variation of the model presents values of 1 in all cases, compared with the original model. And the values of $R_{pos}$ in the original model are below the proposed variation on average 53% for instance 5-5-5-2, 73% for instance 5-5-5-5, 64% for the instances 5-10-10-2, 26% for instance 5-10-15-2, and 18% for instance 5-10-20-2. This indicates that in all cases the direct flow variation between $(i)$ and $(k)$ has always better Pareto fronts as compared with those obtained in the original model.

Concerning the processing time, comparing the original model with the variation that allows the flow from $(i)$ to $(k)$, we have the following: for instance 5-5-5-2 time is increased on average by 1.6% and for instance 5-5-5-5 increased on average by 6.2% over the original model, for instance 5-10-10-2 a decrease is observed on average by 3.4%, and for instances 5-10-15-2 and 5-10-20-2 processing time decreased by 22% and increased by 16%, respectively.

Table 4 shows the results of the variation that allows the flow between distribution centers $(j)$-$(p)$. $D_{avg}$ is greater than 1 in all cases; this indicates a lower quality of the Pareto fronts of this variation compared with the original model.

The $R_{pos}$ values in the original model are on average of 1 with respect to the variation; this indicates that the original model presents always better Pareto fronts. With respect to processing time, for instance 5-5-5-2 time is increased by

Table 4: Results of the third variation.

| Size | Instance | Original model | | Direct flow between ($j$) and ($p$) | | $D_{avg}$ | $D_{min}$ | Time with original model (seconds) | Time flow between ($j$) and ($p$) (seconds) |
|---|---|---|---|---|---|---|---|---|---|
| | | $\lvert S_i \rvert$ | $R_{pos}$ ($S_i$) | $\lvert S_i \rvert$ | $R_{pos}$ ($S_i$) | | | | |
| 5_5_5_2 | 1 | 33 | 1 | 11 | 0.33333333 | 1.0165715 | 1 | 11.631 | 326.195 |
| | 2 | 31 | 1 | 1 | 0.03225806 | 1.01402981 | 1 | 20.896 | 292.527 |
| | 3 | 31 | 1 | 0 | 0 | 1.04683427 | 1.02795867 | 8.728 | 245.921 |
| | 4 | 32 | 1 | 0 | 0 | 1.06735052 | 1.006421 | 13.499 | 1215.217 |
| | 5 | 30 | 1 | 0 | 0 | 1.08324319 | 1.03670392 | 9.841 | 898.683 |
| 5_5_5_5 | 1 | 38 | 1 | 0 | 0 | 1.02172283 | 1.00771899 | 187.27 | 52861.828 |
| | 2 | 39 | 1 | 0 | 0 | 1.01398614 | 1.00201322 | 112.811 | 382637.904 |
| | 3 | 39 | 1 | 6 | 0.15384615 | 1.00778561 | 1 | 3944.193 | 122663.847 |
| | 4 | 40 | 1 | 0 | 0 | 1.01458474 | 1.00033388 | 218.698 | 43666.763 |
| | 5 | 40 | 1 | 2 | 0.05 | 1.01185138 | 1 | 2234.115 | 131359.178 |

average of 460%. For instance 5-5-5-5 time is increased on average by 560% over the original model.

For the analysis of this variation only the first two groups of instances were solved, because the processing time required for evaluating the variation was very long. Solving some instances of the group 5-5-5-5 required up to 34 hours to get results, however for instances 5-10-10-2 after 143 hours of processing the solver did not yield a result, so we decided not testing for the rest of the instances.

## 5. Conclusions

Reference [1] defines the management of the supply chain as the process of planning, implementation, and operational control of the supply chain in an efficient manner. This aspect is defined in the context of tactical decisions that allow for more efficiency in the full cycle of manufacture. The work developed in this research explores an area that has not been sufficiently analyzed and incorporated into mathematical models of supply chain design with selection of distribution channels according to [40, 41].

The CFCLP-TC proposed by [2, 3] incorporates in a novel way the selection of transportation alternatives in the context of a (two-echelon) problem considering plants, distribution centers, and customers. However, the changes proposed in this paper can bring theoretical models to real applications, as they consider situations that could occur in real context.

The first variation that allows the flow to customers from multiple sources (DC) states that the proposed approach generally results in better cost compared with the original proposal where customers are limited to obtaining the product from a single source (DC). This enables better Pareto fronts in the proposed alternative. Although the results obtained suggest that it is cheaper to allow customers to be supplied by more than a distribution center, this increases the complexity of cross-docking, and in accordance with [40] the increase of the level of complexity of the supply chain will negatively affect the performance of a manufacturing plant. However, a greater variety of products may create economies of a scale enough to reduce this effect.

The second variation allows in some cases the product flows from plants directly to customers without necessarily passing through the distribution centers. This variation generally obtained better costs compared to the original proposal. The proposal can be justified when customers are geographically closer to the plants than the distribution centers.

When planning the configuration of the supply chain, it is not possible to accurately predict in the future where they will reside geographically, and it is clear that it is more costly to relocate distribution centers based on the geographical configuration of new customers. Therefore, an alternative that would optimize the cost and time of the supply chain is based on the ability to send directly from plants to customers. The results obtained allowed us to determine the best Pareto fronts with this approach; however, a problem is the time required to solve the instances, since the increase of time is significant as we try to solve the larger instances that resembled real problems. The proposal explores a configuration of the new supply chain that has not been considered in the literature.

The third variation allowing in some cases the exchange of product between distribution centers generally gets worse costs compared with the original model. This proposal is considered in the overall context of managing the supply chain, where a distribution center can supply to another distribution center if the times to get the product to the customer and the costs associated with this alternative are lower. An example is the exchange of product between car dealers to reduce lead times, since placing an order to assembly plant takes time and cost, and these impact the service level.

The results determined that this approach provides Pareto fronts worse than the original model and therefore they have higher costs. For this approach, processing times are greatly increased as the size of the instances grows; this situation is not rare in practice. But one aspect that the model does not consider is the level of customer service, since for the example above it is difficult to predict with accuracy the products requested by the customers, and when this happens, you must deliver the product the customer demands in the shortest possible time and cost.

The proposed approaches are novel and allow mathematical modeling to configure supply chains to situations that occur most commonly in practice and allow closing the gap between theory and practice, contributing to the state of the art in this context.

It is clear that the new approaches produced an increase in the required processing time. The findings and conclusions presented are based only on instances that were tested, which in a context of real implementation are of small size. It is possible that the results of the Pareto front for larger instances may be different; therefore, it is important to determine these fronts, and however the impossibility of doing so with exact methods requires us to try to get them with heuristics and metaheuristics. In the revised literature the Lagrangian relaxation method is widely used to solve similar problems [42], and this may be applied to the problem addressed in this work. Another approach is the use of genetic algorithms proposed by [43] for a similar problem found in the literature. Evolutionary algorithms like Nondominated Sorting Genetic Algorithm (NSGA-II) and Strength Pareto Evolutionary Algorithm (SPEA-II) are widely used in multiobjective problems as in [17, 44, 45].

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The authors appreciate the support of the administration of Polytechnic University of Tulancingo, for the equipment and facilities provided for the development of this work, as well as the comments from the Faculty of Applied Mathematics and Approximation in the same university.

## References

[1] M. T. Melo, S. Nickel, and F. Saldanha-da-Gama, "Facility location and supply chain management—a review," *European Journal of Operational Research*, vol. 196, no. 2, pp. 401–412, 2009.

[2] E. Olivares Benitez, *Capacitated fixed cost facility location problem with transportation choices [Ph.D. thesis]*, ITESM, 2007.

[3] E. Olivares-Benitez, J. L. González-Velarde, and R. Z. Ríos-Mercado, "A supply chain design problem with facility location and bi-objective transportation choices," *TOP*, vol. 20, no. 3, pp. 729–753, 2012.

[4] F. Altiparmak, M. Gen, L. Lin, and T. Paksoy, "A genetic algorithm approach for multi-objective optimization of supply chain networks," *Computers & Industrial Engineering*, vol. 51, no. 1, pp. 196–215, 2006.

[5] A. M. Geoffrion and R. F. Powers, "Twenty years of strategic distribution system design: an evolutionary perspective," *Interfaces*, vol. 25, no. 5, pp. 105–127, 1995.

[6] C. H. Aikens, "Facility location models for distribution planning," *European Journal of Operational Research*, vol. 22, no. 3, pp. 263–279, 1985.

[7] C. S. Revelle and G. Laporte, "The plant location problem: new models and research prospects," *Operations Research*, vol. 44, no. 6, pp. 864–874, 1996.

[8] J. A. Díaz and E. Fernández, "Hybrid scatter search and path relinking for the capacitated $p$-median problem," *European Journal of Operational Research*, vol. 169, no. 2, pp. 570–585, 2006.

[9] N. Mladenović, J. Brimberg, P. Hansen, and J. A. Moreno-Pérez, "The $p$-median problem: a survey of metaheuristic approaches," *European Journal of Operational Research*, vol. 179, no. 3, pp. 927–939, 2007.

[10] M. Baïou and F. Barahona, "On the $p$-median polytope of $Y$-free graphs," *Discrete Optimization*, vol. 5, no. 2, pp. 205–219, 2008.

[11] B. Goldengorin and D. Krushinsky, "Complexity evaluation of benchmark instances for the $p$-median problem," *Mathematical and Computer Modelling*, vol. 53, no. 9-10, pp. 1719–1736, 2011.

[12] M. Sun, "Solving the uncapacitated facility location problem using tabu search," *Computers & Operations Research*, vol. 33, no. 9, pp. 2563–2589, 2006.

[13] I. Averbakh, O. Berman, Z. Drezner, and G. O. Wesolowsky, "The uncapacitated facility location problem with demand-dependent setup and service costs and customer-choice allocation," *European Journal of Operational Research*, vol. 179, no. 3, pp. 956–967, 2007.

[14] H.-C. Huang and R. Li, "A $k$-product uncapacitated facility location problem," *European Journal of Operational Research*, vol. 185, no. 2, pp. 552–562, 2008.

[15] A. Klose and A. Drexl, "Facility location models for distribution system design," *European Journal of Operational Research*, vol. 162, no. 1, pp. 4–29, 2005.

[16] C.-H. Chen and C.-J. Ting, "Combining Lagrangian heuristic and ant colony system to solve the single source capacitated facility location problem," *Transportation Research E*, vol. 44, no. 6, pp. 1099–1122, 2008.

[17] L. Lin, M. Gen, and X. Wang, "Integrated multistage logistics network design by using hybrid evolutionary algorithm," *Computers & Industrial Engineering*, vol. 56, no. 3, pp. 854–873, 2009.

[18] Z. Yao, L. H. Lee, W. Jaruphongsa, V. Tan, and C. F. Hui, "Multi-source facility location-allocation and inventory problem," *European Journal of Operational Research*, vol. 207, no. 2, pp. 750–762, 2010.

[19] İ. K. Altınel, E. Durmaz, N. Aras, and K. C. Özkısacık, "A location-allocation heuristic for the capacitated multi-facility Weber problem with probabilistic customer locations," *European Journal of Operational Research*, vol. 198, no. 3, pp. 790–799, 2009.

[20] K. E. Caggiano, P. L. Jackson, J. A. Muckstadt, and J. A. Rappold, "Efficient computation of time-based customer service levels in a multi-item, multi-echelon supply chain: a practical approach for inventory optimization," *European Journal of Operational Research*, vol. 199, no. 3, pp. 744–749, 2009.

[21] A. Klose and S. Görtz, "A branch-and-price algorithm for the capacitated facility location problem," *European Journal of Operational Research*, vol. 179, no. 3, pp. 1109–1125, 2007.

[22] A. A. Kuehn and M. J. Hamburger, "A heuristic program for locating warehouses," *Management Science*, vol. 9, no. 4, pp. 643–666, 1963.

[23] B. M. Khumawala, "An efficient heuristic procedure for the capacitated warehouse location problem," *Naval Research Logistics Quarterly*, vol. 21, no. 4, pp. 609–623, 1974.

[24] M. R. Korupolu, C. G. Plaxton, and R. Rajaraman, "Analysis of a local search heuristic for facility location problems," in *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '98)*, pp. 1–10, ACM, Philadelphia, Pa, USA, 1998.

[25] E. Rolland, D. A. Schilling, and J. R. Current, "An efficient tabu search procedure for the $p$-median problem," *European Journal of Operational Research*, vol. 96, no. 2, pp. 329–342, 1997.

[26] H. Delmaire, J. A. Díaz, E. Fernández, and M. Ortega, "Reactive GRASP and tabu search based heuristics for the single source capacitated plant location problem," *INFOR Journal*, vol. 37, no. 3, pp. 194–225, 1999.

[27] T. L. Magnanti and R. T. Wong, "Accelerating benders decomposition: algorithmic enhancement and model selection criteria," *Operations Research*, vol. 29, no. 3, pp. 464–484, 1981.

[28] J. M. Y. Leung and T. L. Magnanti, "Valid inequalities and facets of the capacitated plant location problem," *Mathematical Programming*, vol. 44, no. 1–3, pp. 271–291, 1989.

[29] K. Aardal, "Capacitated facility location: separation algorithms and computational experience," *Mathematical Programming*, vol. 81, no. 2, pp. 149–175, 1998.

[30] G. Laporte, F. V. Louveaux, and L. van Hamme, "Exact solution of a stochastic location problem by an integer L-shaped algorithm," *Transportation Science*, vol. 28, no. 2, pp. 95–103, 1994.

[31] S. Tragantalerngsak, J. Holt, and M. Rönnqvist, "Lagrangian heuristics for the two-echelon, single-source, capacitated facility location problem," *European Journal of Operational Research*, vol. 102, no. 3, pp. 611–625, 1997.

[32] L. A. Moncayo-Martínez and Z. Z. David, "Optimising safety stock placement and lead time in a assembly supply chain using bi-objetive MAX-MIN ant system," *International Journal of Production Economics*, vol. 145, no. 1, pp. 18–28, 2013.

[33] W.-C. Yeh and M.-C. Chuang, "Using multi-objective genetic algorithm for partner selection in green supply chain problems," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4244–4253, 2011.

[34] R. Rajesh, S. Pugazhendhi, and K. Ganesh, "Simulated annealing algorithm for balanced allocation problem," *The International Journal of Advanced Manufacturing Technology*, vol. 61, no. 5–8, pp. 431–440, 2011.

[35] G. C. Cabrera Guillermo, E. Cabrera, R. Soto, L. J. M. Rubio, B. Crawford, and F. Paredes, "A hybrid approach using an artificial bee algorithm with mixed integer programming applied to a large-scale capacitated facility location problem," *Mathematical Problems in Engineering*, vol. 2012, Article ID 954249, 14 pages, 2012.

[36] Y.-S. Myung, H.-G. Kim, and D.-W. Tcha, "A bi-objective uncapacitated facility location problem," *European Journal of Operational Research*, vol. 100, no. 3, pp. 608–616, 1997.

[37] J. G. Villegas, F. Palacios, and A. L. Medaglia, "Solution methods for the bi-objective (cost-coverage) unconstrained facility location problem with an illustrative example," *Annals of Operations Research*, vol. 147, no. 1, pp. 109–141, 2006.

[38] R. D. Galvão, L. G. A. Espejo, B. Boffey, and D. Yates, "Load balancing and capacity constraints in a hierarchical location model," *European Journal of Operational Research*, vol. 172, no. 2, pp. 631–646, 2006.

[39] M. da Graça Costa, M. E. Captivo, and J. Clímaco, "Capacitated single allocation hub location problem—a bi-criteria approach," *Computers & Operations Research*, vol. 35, no. 11, pp. 3671–3695, 2008.

[40] C. C. Bozarth, D. P. Warsing, B. B. Flynn, and E. J. Flynn, "The impact of supply chain complexity on manufacturing plant performance," *Journal of Operations Management*, vol. 27, no. 1, pp. 78–93, 2009.

[41] J. Mula, D. Peidro, M. Díaz-Madroñero, and E. Vicens, "Mathematical programming models for supply chain production and transport planning," *European Journal of Operational Research*, vol. 204, no. 3, pp. 377–390, 2010.

[42] H. Lidestam and M. Rönnqvist, "Use of Lagrangian decomposition in supply chain planning," *Mathematical and Computer Modelling*, vol. 54, no. 9-10, pp. 2428–2442, 2011.

[43] M. Gen, F. Altiparmak, and L. Lin, "A genetic algorithm for two-stage transportation problem using priority-based encoding," *OR Spectrum*, vol. 28, no. 3, pp. 337–354, 2006.

[44] X. Li, L. Amodeo, F. Yaloui, and H. Chehade, "A multiobjetive optimization approach to solve parallel machines scheduling problem," *Advances in Artificial Intelligence*, vol. 2010, Article ID 943050, 10 pages, 2010.

[45] H. Xing and R. Qu, "A nondominated sorting genetic algorithm for bi-objetive network coding based multicast routing problems," *Information Sciences*, vol. 233, pp. 36–53, 2013.