International Journal of
Operations and Quantitative Management

Lappas, Kritikos, Ioannou
Evolutionary Optimization Algorithm for the Inventory Routing Problem

Wang, Chou
Combination of Renewable Energies for an Enterprise

Ratnayake, Perera
Incentive Model for Operational Level Employee Teams

Woo, Kim, Lim
Data Based Decision Making for the Selection of Celebrity Advertising Model

Volume 24 Number 2 June 2018

International Journal of
Operations and Quantitative Management

AIMS AND SCOPE

The objective of the International Journal of Operations and Quantitative Management (IJOQM) is to provide an international forum for discussion of advancements in operations management, operations research, quantitative management and management science. The Journal publishes high-quality research articles in all these areas. The Journal particularly encourages articles dealing with international issues, practical experiences, pedagogical aspects including case research, developments in the emerging nations, service management, and other nontraditional and innovative areas. Purely theoretical articles should make significant contribution to the existing knowledge.

Editor-in-Chief

Omprakash K. Gupta	*Phone: 1-713-222-5343*
University of Houston-Downtown	*Fax: 1-713-226-5226*
College of Business	*e-mail: editor@ijoqm.org*
Houston, TX 77002, USA	*website: www.ijoqm.org*

The International Journal of Operations and Quantitative Management (ISSN 1082-1910) is published quarterly (March, June, September and December) by International Forum of Management Scholars. Subscriptions, changes of address, and all correspondence should be sent to: 12436 FM 1960 W. Pmb #140, Houston, TX 77065.

International Journal of
Operations and Quantitative Management

Volume 24, Number 2, June 2018

Contents

A Hybrid Evolutionary Optimization Algorithm for the Inventory Routing Problem 75
Pantelis Z. Lappas, Manolis N. Kritikos, George D. Ioannou

Optimal Combination of Renewable Energies for an Enterprise–A Case of Taiwan 117
Hsiao-Fan Wang, Pei-Wen Chou

Incentive Model for Operational Level Employee Teams in an Organization 139
Vijitha Ratnayake, Gayathri Perera

Big data Based Decision Making for the Selection of Celebrity Advertising Model: A Korean Case 153
Mina Woo, Jinwha Kim, Gyoo Gun Lim

EDITORIAL REVIEW BOARD

JOURNAL COSPONSORS

A Hybrid Evolutionary Optimization Algorithm for the Inventory Routing Problem

Pantelis Z. Lappas
Manolis N. Kritikos
George D. Ioannou
Athens University of Economics and Business
(pzlappas@aueb.gr)
(kmn@aueb.gr)
(ioannou@aueb.gr)

Volume 24, *Number* 2
June 2018, *pp.* 75-115

The main objective of this paper is to propose a hybrid evolutionary optimization algorithm for solving the Inventory Routing Problem (IRP). The IRP arises from the application of the Vendor Managed Inventory (VMI) concept, where the supplier (vendor) has to make inventory and routing decisions simultaneously for a given planning horizon. This paper focuses on a scenario where a single-product type has to be delivered by a fleet of capacitated homogenous vehicles and housed at a depot over a finite and discrete planning horizon. The demand is fully available to the decision maker (supplier) at the beginning of the planning horizon, stock-outs are not allowed, and transportation costs and inventory holding costs of customers are taken into account in the objective function. Due to the NP-hard nature of the IRP, it is very difficult to develop an exact algorithm that can solve large-scale problems within a reasonable computation time. As an alternative, a hybrid evolutionary optimization algorithm based on two well-known meta-heuristics, the Genetic Algorithm and the Simulated Annealing Algorithm, is presented to handle the IRP. Namely, the Genetic Algorithm is related to the planning phase, while the Simulated Annealing Algorithm is associated with the routing phase. A repetitive procedure, containing characteristics from both referred meta-heuristics, is applied to obtain a near-optimal feasible solution. Testing instances with different properties are established to investigate algorithmic performance, and the computational results are then reported.

Keywords: Routing, Inventory Routing Problem, Genetic Algorithm, Simulated Annealing, Evolutionary Optimization

1. Introduction

In recent years, the Inventory Routing Problem (IRP) has received a great deal of attention from academics, consultants and practitioners. It reflects a multi-functional problem that attempts to integrate two different functions within the supply chain network, i.e., planning and routing (Min and Zhou, 2002). In particular, planning is associated with the Inventory Control Problem (ICP), while routing is related to the Vehicle Routing Problem (VRP). The ICP represents an activity that aims to organize the availability of goods to customers during a given planning horizon (Axsäter, 2006), while the VRP concerns the distribution of goods between suppliers and customers, without taking into account the time scope (Toth and Vigo, 2002).

Whereas VRPs typically deal with a single period (e.g., a day), IRPs have to deal with a longer horizon (multi periods: e.g., a sequence of days).

In the context of the IRP, these two widely studied problems in the Operations Research literature are modeled simultaneously since an inter-relationship exists between them (Moin and Salhi, 2007; Archetti and Speranza, 2016). If only the ICP for the customers is concerned and the VRP for the supplier is ignored, the supply chain cost, including the total transportation and total inventory cost, is not minimized optimally, as the VRP decisions cannot be made effectively and vice versa. The IRP arises in environments where Vendor Managed Inventory (VMI) policies are applied. It can be assumed as an extension of the VRP, which integrates routing and inventory allocation decisions. Analytically, the vendor (supplier) monitors the inventory levels of the customers and determines (a) the delivery times (when to visit his customers), (b) the quantities (how much to deliver to each of them when they are served), so that stock-outs are avoided, and (c) the set of routes used by a fleet of vehicles to serve a given set of customers (how to integrate the customers into the vehicle routes).

IRPs can be categorized into three levels (Andersson et al., 2010; Coelho et al., 2013). The first categorization is based on the structural variants presented in IRPs, namely, product, time horizon, network topology, routing, inventory policy, inventory decisions, fleet composition and fleet size. The second categorization is related to the availability of information on the demand, reflecting several types of IRPs, for example, deterministic, stochastic, and dynamic and stochastic IRPs. Moreover, the third categorization is associated with the chosen solution approach. According to Ballou (1989) the modeling of supply chain and logistics problems has traditionally relied on three primary methods, i.e., simulation, optimization (exact algorithm) and heuristics, which can be divided into two categories (Griffis et al., 2012): classic heuristics (construction heuristics, local improvement heuristics) and meta-heuristics (local search meta-heuristics and population search meta-heuristics). The recent literature has shown an increased interest in so-called matheuristics, methods that combine exact and heuristic approaches (Maniezzo et al., 2009). Archetti and Speranza (2013) classified matheuristics into three classes: decomposition approaches, improvement heuristics and column generation-based approaches.

It is worth noting that IRP decisions can be (a) decisions over time only, in which the delivery times and the quantities have to be determined at the same time, while the routes are given, and (b) decisions over time and space, where delivery times, quantities and routes have to be determined simultaneously (Bertazzi and Speranza, 2012; Bertazzi and Speranza, 2013). Furthermore, the optimal solution of an IRP depends on the objective function that has been chosen (Bertazzi et al., 2008). As a result, an objective function can be (a) the sum of transportation costs only, (b) the sum of transportation and inventory holding costs of the customers or (c) the sum of transportation and inventory holding costs of the supplier and the customers. It should not pass unnoticed that under the VMI concept, stock-outs are not allowed, and therefore, the objective function does not include shortage costs. In general, the objective function of IRP reflects a minimum-constrained problem, broadly studied in the literature (e.g., Neubert et al., 2010; Savino et al., 2014; Savino and Mazza, 2015).

In this paper, the main objective is to propose an approach for solving the IRP with the following characteristics. A single-product type has to be delivered by a fleet of capacitated homogenous vehicles (multiple vehicles) housed at a depot over a finite and discrete planning horizon. The network topology taken into account by the IRP model is one-to-many; that is, one supplier serves many geographically dispersed customers (demand points). A vehicle can visit more than one customer (multiple routing), while a vehicle's trip starts and ends at the depot (supplier). As far as the inventory policy is concerned, a Maximum Level (ML) policy is considered, in which any customer has defined a maximum inventory level and every time a customer is served, the delivered quantity is such that the inventory level at the customer is not greater than the maximum level. It is assumed that the depot has a sufficient supply of products that can cover all customers' demands throughout the planning horizon. Moreover, the inventory is not allowed to become negative (fixed inventory) since the lowest inventory level is either fixed or equal to zero. With respect to the availability of information on customer demand, the proposed IRP model is deterministic since the demand is fully available to the supplier at the beginning of the planning horizon.

Regarding the solution approach, a hybrid evolutionary optimization algorithm that combines a nature-inspired optimization algorithm (local search meta-heuristic), such as the Simulated Annealing Algorithm (SA), as well as a biologically-inspired optimization algorithm (population search meta-heuristic), that is, the Genetic Algorithm (GA), is presented to handle the problem. The SA is associated with the routing decisions (routing phase), while GA is related to the inventory allocation decisions (planning phase). A repetitive procedure, containing characteristics of both meta-heuristics, is applied to obtain a near-optimal feasible solution. In addition, IRP decisions are decisions over time and space, while the objective function represents the sum of transportation and inventory holding costs of the customers.

The remainder of the paper is organized as follows. Section 2 presents an overview of the state of the art in research on the Inventory Routing Problem. A problem description and mathematical formulation are presented in Section 3. The proposed solution approach is described and analyzed in detail in Section 4. Section 5 presents computational results, while in Section 6, conclusions and future directions are given.

2. State of the Art

Routing problems have attracted attention as a possible solution to many of the complex issues surrounding Supply Chain Management (ScM). In today's economic environment, efficiency for firms is moving from an internal to a supply chain priority since the competition is not among them, but among their supply chains (Croom et al., 2000; Tan, 2001). As a consequence, the ultimate success of a firm depends on its ability to integrate and coordinate different supply chain activities within the supply chain network (Min and Zhou, 2002; Schmid et al., 2013). Routing problem (RP) is the generic name given to a whole class of problems in which transportation is necessary (Diaz-Parra et al., 2014). The issue of RPs can be addressed in two dimensions: (a) classical routing problems, such as the Traveling Salesman Problem (TSP) and the Vehicle Routing Problem (VRP), and (b) highly relevant extensions of classical routing problems like the Inventory Routing Problem (IRP) and the Production Routing Problem (PRP).

The TSP is the most basic routing problem and a typical model of the combinatorial optimization problems whose computation complexity is derived from non-polynomial time (NP-hard problem). In particular, the problem is to find the shortest route (minimum transportation cost) that starts from a depot, visits all customers exactly once, and returns to the depot (Flood, 1956). For a comprehensive review of the proposed solution approaches including exact Algorithm, heuristics and meta-heuristics, see Laporte (2010), Rego et al. (2011) and Arram et al. (2014). However, in transportation problems, customers usually have a demand, whereas the depot consists of a fleet of vehicles with limited and known capacity. This situation reflects the VRP (Dantzig and Ramser, 1959), which generalizes the Multiple Traveling Salesman Problem (m-TSP), i.e., the TSP with m vehicles (Bektas, 2006). A survey of the VRP literature as well as the most important exact solutions, classical and modern heuristics are presented by Cordeau et al. (2002), Eksioglu et al. (2009), Laporte (2009) and Potvin (2009). The Vehicle Routing Problem with Time Windows (VRPTW) is a generalization of the VRP involving the added complexity that every customer should be served within a given time window (Bräysy and Gendreau, 2005a; Bräysy and Gendreau, 2005b; El-Sherbeny, 2010).

Furthermore, the IRP is an extension of the VRP, which integrates routing decisions with inventory control (Moin and Salhi, 2007; Andresson et al., 2010; Coelho et al., 2013; Archetti and Speranza, 2016). The problem arises in environments where VMI policies are employed, while the supplier decides the delivery times, the quantities and the vehicle routes at the same time. The main objective is to minimize the total transportation and inventory holding costs. From the perspective of the design of multi-echelon distribution networks, the work of Melachrinoudis et al. (2009) could be mentioned due to the usual one-to-many network topology of the IRPs. They presented a Voronoi diagram to design a two-echelon distribution network. The proposed Voronoi diagram can systematically reduce the number of warehouse-customer assignments by exploiting its proximity relationships. Other works worth mentioning focus on the stochastic variant of the multi-echelon distribution networks. For instance, Nikzad et al. (2017) described a risk pooling strategy in order to reduce variability by aggregating customer demand across products, time, and location. Since order acceptance/rejection decisions are often decisions under uncertainty (demand uncertainty), Sujan et al. (2015) proposed a method to simultaneously quote the due date and the price of each incoming order when the contingent orders exist.

The Inventory Routing Problem with Hard or Soft Time Windows (IRPTW/IRPSTW), which has not been excessively researched in the literature, is a generalization of the standard IRP involving the added complexity that every customer should be served within a given time window. Liu and Lee (2011) proposed a two-phase heuristic method for solving the IRPSTW. The first phase of the heuristic algorithm finds an initial solution based on a construction approach, while the second phase improves the initial solution by adopting a variable neighborhood tabu search algorithm. In addition, Zeng and Zhao (2010) represented the stochastic IRPSTW as a discrete time Markov decision process model and solved it by using dynamic programming approximations. Lappas et al. (2015a) presented a two-phase solution algorithm based on the Monte Carlo Simulation and the Genetic

Algorithm to solve the IRPTW. The first phase is related to the planning phase of the IRPTW, in which delivery times and quantities are determined by implementing the well-known inventory policy (s, S) for inventory management using the Monte Carlo Simulation. In the second phase, the Genetic Algorithm is applied to combine the customers into the vehicle routes by solving a VRPTW for a specific time period during the planning horizon. Some applications in the context of IRPTW/IRPSTW were presented by Zhang et al. (2013), Li et al. (2015) and Zhang et al. (2015). The IRPTW is obviously NP-hard, being a generalization of the IRP, which reduces to the TSP when the planning horizon is equal to a single period (e.g., one day); there are no inventory holding costs; all the customers need to be served but not in specific time windows; there is a single vehicle and transportation capacity is infinite (Bertazzi and Speranza, 2013; Lappas et al., 2015b; Lappas et al., 2015c) (Figure 1).

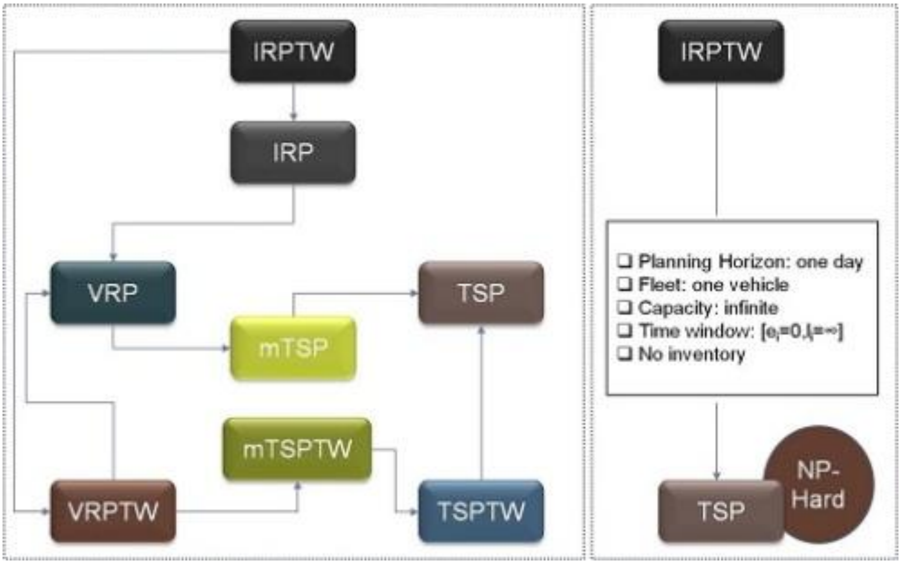


Figure 1 *NP-Hard Nature of the IRP*

The PRP is also a core problem that has to be solved specifically in a VMI replenishment system and can be assumed to be the generalization of the IRP. The vendor monitors the inventory levels of the customers, while production, inventory, distribution and routing decisions have to be made simultaneously. Actually, PRP combines the Lot-sizing Problem and the VRP. A large body of literature on PRP can be found. Chen and Vairaktarakis (2005) studied an integrated scheduling model of production and distribution operations. Liu et al. (2007) presented a solution approach to solve the integrated production and distribution problem in the context of a real chemical supply chain in North America, while Chen et al. (2009) proposed a nonlinear model to consider production scheduling and vehicle routing with time windows for perishable products. In addition, Savino et al. (2014) provided a very interesting approach to study a multiple-objective flow-shop modeling and scheduling problem by multi-agent system. Several formulations and Branch-and-Cut Algorithm for multi-vehicle PRP were proposed by Adulyasak et al. (2014).

Moreover, heuristics (e.g., Mabrouk et al., 2010) and meta-heuristic (e.g., Staggemeier et al., 2002) approaches were studied. For a comprehensive review of this literature through the year 2015, see Chen (2004) and Adulyasak (2015).

Several applications of the IRP have been found. The result of an analysis of the scientific literature led to the identification of six main paths of development in the overall field of the IRP: (1) maritime transportation (Ronen, 1993; Arga et al., 2013; Song and Furman, 2013; Hewitt et al., 2013; Arga et al., 2014; Papageorgiou et al., 2014; Arga et al., 2015; Jiang and Grossmann, 2015; Hemmati et al., 2015; Arga et al., 2016a; Arga et al., 2016b; Hemmati et al., 2016), (2) industrial gas distribution (Bell et al., 1983; Goel et al., 2012; Ghiami et al., 2015; Shao et al., 2015; Singh et al., 2015; Goel et al., 2015; Andersson et al., 2016), (3) distribution of perishable goods (Federgruen and Zipkin, 1984; Federgruen et al., 1986; Le et al., 2013; Soysal et al., 2015; Mirzaei and Seifi, 2015; Soysal et al., 2016; Diabat et al., 2016), (4) fuel delivery (Popović et al., 2012), (5) medical waste collection (Nolz et al., 2014a; Nolz et al., 2014b) and medical drug distribution (Niakan and Rahimi, 2015), in addition to (6) distribution of agriculture products (Liao et al., 2013) and groceries (Mercer and Tao, 1996; Gaur and Fisher, 2004).

The IRP research can be divided into three main streams. In the first stream, exact Algorithm have been proposed to solve the IRP. Some of the exact Algorithm that have been published through the year 2013 and that can solve an IRP are summarized by Coelho et al. (2013) and Coelho and Laporte (2013). The second stream of research contains approximation approaches. Due to the inability of the exact Algorithm to solve large-scale IRP instances, an impressive number of heuristics as well as meta-heuristics have been proposed. Constructive heuristics and improvement heuristics have been developed and presented by Abdelmaguid et al. (2009) for the IRP with backlogging. The proposed construction heuristic, called ETCH (Estimated Transportation Costs Heuristic), estimates a transportation cost value for each customer in each time period to facilitate a comparison between the transportation and the inventory holding and shortage costs. Due to the myopic nature of the ETCH and the fact that partial fulfillment of demand is not allowed, an improvement heuristic was proposed in order to overcome the above limitations. The improvement heuristic is based on the idea of exchanging customer delivery quantities between periods to allow transitions from a given solution to its neighborhood. More recently, Raa (2015) provided a multi-start two-phase heuristic solution method consisting of an insertion-based construction phase and an improvement phase for the Cyclic IRP, while Nambirajan et al. (2016) proposed a three-phase heuristic called CARE (Clustering, Allocation, Routing, Extended) for two-stage multi-product inventory routing problems with replenishments.

Furthermore, several local search meta-heuristics such as Tabu Search (TS) (Archetti et al., 2012; Li et al., 2014; Qin et al., 2014), Greedy Randomized Adaptive Search Procedure (GRASP) (Guemri et al., 2016), Iterated Local Search (ILS) (Vansteenwegen and Mateo, 2014; Santos et al., 2016), Variable Neighborhood Search (VNS) (Mjirda et al., 2012; Mjirda et al., 2014, Mjirda et al., 2016) and Adaptive Large Neighborhood Search (ALNS) (Coelho et al., 2012a; Aksen et al., 2014; Shirokikh and Zakharov, 2015) have been applied to the IRP. An alternative approach that combines simulation with heuristics has been presented by Juan et al. (2014), who described and used a “simheuristic” algorithm to solve the single-period stochastic IRP with stock-outs. Their approach combines the Monde Carlo Simulation with the multi-start randomized heuristic.

A number of population search meta-heuristics have been proposed for the solution of the IRP and its variants. Huang and Lin (2010) presented a modified ant colony optimization algorithm for multi-item IRPs with demand uncertainty. Tatsis et al. (2013) described the multiple suppliers, one retailer (many-to-one) IRP and proposed an ant-based optimization algorithm to solve the problem. In both papers, the main objective is to minimize the total transportation, inventory holding and backlogging costs. A hybrid heuristic method that integrates a Large Neighborhood Search (LNS) into Particle Swarm Optimization (PSO) presented by Liu et al. (2015) to solve the Periodic IRP. In addition, Yang et al. (2015) applied indicator-based evolutionary Algorithm and swarm Algorithm to find an approximation to the Pareto front of the IRP. Evolutionary optimization Algorithm, such as GAs, have also been proposed to solve the IRP. This is particularly clear in the studies cited by Abdelmaguid and Dessouky (2006), Aziz and Moin (2007), Moin et al. (2011) Simić and Simić (2013), Shukla et al. (2013), Cho et al. (2013) and Park et al. (2016).

The third stream of research is associated with mathheuristics, consisting of decomposition approaches (Campbell and Savelsbergh, 2004), improvement heuristics (Coelho et al., 2012b; Bertazzi et al., 2013; Guerrero et al., 2013; Archetti et al., 2014; Bertazzi et al., 2015) and column generation-based approaches (Aghezzaf et al., 2006).

The research presented below represents an attempt to use local search and population search meta-heuristics to solve the IRP. The basic idea of the proposed approach is to combine a nature-inspired evolutionary optimization algorithm, such as the SA, and a biologically-inspired evolutionary optimization algorithm, that is, the GA, to handle the IRP. Therefore, a hybrid evolutionary optimization algorithm is proposed to solve the IRP. The SA is associated with the routing phase of the IRP, while the GA is related to the planning phase of the IRP. Both Algorithm are dealt with in an iterative way.

The works most closely related to this paper are most likely those of Abdelmaguid and Dessouky (2006), Aziz and Moin (2007), Moin et al. (2011), Cho et al. (2013), and Park et al. (2016). Abdelmaguid and Dessouky (2006) introduced a genetic algorithm to solve the one-to-many type of the IRP with finite horizon. The objective function includes transportation costs as well as inventory holding and shortage costs on the end inventory positions. In particular, they designed a genetic representation in the form of a two-dimensional matrix based on the delivery schedule and addressed the vehicle routing part using the Clarke and Wright algorithm. In addition, a randomized version of a construction heuristic called ATCH (Approximate Transportation Costs Heuristic) was used to generate the initial random population, while suitable crossover and mutation operators were designed for the improvement phase of the genetic algorithm. In the studies by Aziz and Moin (2007) and Moin et al. (2011), the many-to-one type of IRP with finite horizon is addressed. Both transportation and inventory costs are considered, while a hybrid genetic algorithm combining a genetic algorithm (planning phase) and a simple 2-opt procedure (routing phase) is presented. Cho et al. (2013) proposed an adaptive genetic algorithm for the time dependent inventory routing problem considering the one-to-many network topology. This paper takes into account the effect of dynamic traffic conditions in an urban context, while the objective function consists of the transportation, inventory holding and shortage costs at the end of the period inventory positions. More recently, Park et al. (2016) presented a genetic algorithm for the inventory routing problem with lost sales under a VMI strategy in a two-echelon supply chain comprised of a single manufacturer and multiple retailers (one-to-many network topology). The objective function consists of the transportation costs, the inventory holding cost of the manufacturer, the inventory holding costs of the retailers and the costs associated with lost sales.

Most of the previous research has considered a one-to-many type of IRP in which the objective function includes shortage costs at the end of the period inventory positions. In addition, some of the previously reported research (e.g., Abdelmaguid and Dessouky, 2006; Aziz and Moin, 2007; Moin et al., 2011) has focused only on the planning phase of the IRP, while the routing phase has been addressed by simple heuristics such as the Clarke and Wright algorithm and the 2-opt algorithm. In this study, stock-outs are not allowed, while an emphasis is given to how a population-based search meta-heuristic for the planning phase can be used in hybrid synthesis with a single-point search meta-heuristic for the routing phase of the problem. The computational study demonstrates the effectiveness of the proposed algorithm and underscores the importance of integrating the inventory and vehicle routing decisions. Finally, this paper provides various graphical presentation formats to highlight the insights that are gained. In particular, the analytical results and graphic presentations help to simplify complicated issues and convey meaningful insights into the problem.

3. Problem Description and Mathematical Formulation

This section presents a modeling framework for formulating the IRP.Let $G = (V, E)$ be a complete undirected graph where $V = \{0, \dots, n\}$ is the set of vertices and $E = \{(i, j): i, j \in V, j > i\}$ is the set of edges. Vertices $1, \dots, n$ correspond to the customers, whereas vertex 0 corresponds to the supplier (depot). The model presented here deals with the repeated distribution of a single product from a single supplier to a set of geographically dispersed customers $C = V \backslash \{0\} = \{1, \dots, n\}$ over a given time horizon of length H. The set of time horizons is denoted by$T = \{1, \dots, H\}$. Each customer $i \in C$ faces a different demand d_i^t per time period $t \in T$, maintains his own inventory up to capacity U_i, and incurs an inventory holding cost of h_i per period per unit. It is assumed that the depot has a sufficient supply of items that can cover all customers' demands throughout the planning horizon, that is, $U_0 = +\infty$.

A nonnegative cost, c_{ij} is associated with each edge $(i, j) \in E$ and represents the travel cost spent to go from vertex i to vertex j $\forall i, j \in V$. Generally, the usage of the loop edge, (i, i) is not allowed, and this is imposed by defining $c_{ii} = +\infty$ for all $i \in V$. In addition, the cost matrix satisfies the triangle inequality:$c_{ik} + c_{kj} \geq c_{ij}$. In other words, it is not convenient to deviate from the direct link between two vertices. Since G is a complete undirected graph, the cost matrix $[c_{ij}]$ is symmetric, and as a result, $c_{ij} = c_{ji}\ \forall\, i, j \in V$. Vertices are associated with points of the plane having the given coordinates $(x_i, y_i) \forall i \in V$. The cost c_{ij} for each edge $(i, j) \in E$ is defined as the Euclidean distance between the two vertices $i, j \in V$. Therefore, $c_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$.

An unlimited fleet of identical vehicles with capacity Q is available for the distribution of the product. The fleet of vehicles is denoted by the set $K = \{1,2,\ldots\}$.However, to model the problem, an upper bound on the number of vehicles needed to distribute the products should be defined. A trivial upper bound on the maximum fleet size needed is $|K| = |C| = n$. A set of vehicle routes (one route for each vehicle) for the fleet of vehicles, based on the depot, must be determined for a set of customers in each time period$t \in T$. Furthermore, the formulation uses the following decision variables

- w_{ik}^t: The amount of delivery to customer $i \in C$ in period $t \in T$participating in vehicle route$k \in K$.
- x_{ijk}^t: The number of times the edge $(i,j) \in E$ is traversed by vehicle $k \in K$ in period $t \in T$. Actually, this variable is a binary variable, equal to 1 if and only if vehicle $k \in K$ traverses edge $(i,j) \in E$ in period $t \in T$.
- y_{ik}^t: A binary variable that is used to assign customers to vehicles, with value 1 indicating that customer $i \in C$ will be visited by vehicle $k \in K$ in period $t \in T$, and 0 otherwise.
- I_i^t: A nonnegative variable indicating the inventory level at customer $i \in C$ at the end of period $t \in T$. It should be mentioned that at the beginning of the planning horizon, each customer $i \in C$ has an initial inventory level of $I_i^0 = 0, \forall i \in C$ of product.

Moreover, stock-outs are not allowed at the customers, while the quantities delivered by each vehicle in each route cannot exceed the vehicle capacity. As far as the replenishment policy is concerned, a Maximum Level (ML) policy is applied. Therefore, any customer has defined a maximum inventory level. Every time a customer is served, the delivered quantity is such that the inventory level at the customer is not greater than the maximum level. After defining the necessary parameters and decision variables, the IRP can be formulated as a mixed integer linear programming as shown below.

$\text{Min}\left(\sum_{k\in K}\sum_{t\in T}\sum_{i\in V}\sum_{j\in V,j>i} c_{ij}x_{ijk}^t + \sum_{t\in T}\sum_{i\in C} h_i I_i^t\right)$	(1)
Subject to:	
$I_i^t = I_i^{t-1} + \sum_{k\in K} w_{ik}^t - d_i^t, \forall i \in C, \forall t \in T$	(2)
$I_i^t \geq 0, \forall i \in C, \forall t \in T$	(3)
$\sum_{k\in K} w_{ik}^t + I_i^{t-1} \leq U_i, \forall i \in C, \forall t \in T, \forall k \in K$	(4)
$\sum_{i\in C} w_{ik}^t \leq Q y_{ik}^t, \forall k \in K, \forall t \in T$	(5)
$w_{ik}^t \leq U_i y_{ik}^t, \forall i \in C, \forall k \in K, \forall t \in T$	(6)

$\sum_{k\in K} y_{ik}^t \le 1, \forall i \in C, \forall t \in T$	(7)
$\sum_{j\in V, j>i} x_{ijk}^t + \sum_{j\in V, j<i} x_{jik}^t = 2y_{ik}^t, \forall i \in C, \forall k \in K, \forall t \in T$	(8)
$\sum_{i\in S} \sum_{j\in S, j>i} x_{ijk}^t \le \sum_{i\in S} y_{ik}^t - y_{sk}^t, \forall S \subseteq C, \forall s \in S, \forall k \in K, \forall t \in T$	(9)
$w_{ik}^t \ge 0, \forall i \in C, \forall k \in K, \forall t \in T$	(10)
$x_{ijk}^t \in \{0,1\}, \forall i \in C, \forall j \in C, j > i, \forall k \in K, \forall t \in T$	(11)
$x_{0jk}^t \in \{0,1,2\}, \forall j \in C, \forall k \in K, \forall t \in T$	(12)
$y_{ik}^t \in \{0,1\}, \forall i \in V, \forall k \in K, \forall t \in T$	(13)

The objective function (1) minimizes the total transportation and inventory costs. The transportation cost $c_{ij}, \forall i,j \in V$ is taken into account only if a vehicle $k \in K$ traverses edge $(i,j) \in E$ in period $t \in T$ (i.e, $x_{ijk}^t = 1$). The inventory holding cost $h_i, \forall i \in C$ takes place if only an inventory level in period $t \in T$ exists (i.e., $I_i^t > 0$). Constraints (2) are the inventory balance equations for all the customers. For customer $i \in C$, they define the inventory level at$i$at the end of period $t \in T$ by its inventory level at the end of period $t - 1$ plus the amount of delivery to customer i in period t, using a potential vehicle route $k \in K$, minus the demand of customer i in period t. Constraints (3) guarantee that no stock-out occurs at any customer $i \in C$ during the planning horizon, and constraints (4) limit the inventory level of the customers to the corresponding maximum inventory level (ML policy). Constraints (5) ensure that the vehicle capacities are not exceeded in any period $t \in T$ during the planning horizon.Namely, the total quantity delivered by a vehicle $k \in K$in period $t \in T$to a set of customers cannot be greater than its capacity$Q$. Constraints (6) impose the condition that if customer $i \in C$ is visited in period $t \in T$, any quantity is delivered to the customer $i$is limited to the customer's inventory capacity, and this bound is tightened by constraints (4). As a consequence, it is impossible to deliver to customer $i \in C$ quantity more than U_i. In addition, constraints (7) guarantee that a customer $i \in C$can be visited exactly once in each period $t \in T$. Therefore, in each period $t \in T$, a customer $i \in C$ is associated only with one vehicle $k \in K$, i.e., he belongs only to one vehicle route. Constraints (5), (8) and (9) are the routing constraints. They guarantee that feasible routes are determined to visit all customers served in period $t \in T$. Analytically, constraints (8) ensure that if deliveries are made in period $t \in T$, then the vehicle route travelled in period t has to contain one edge entering every vertex i of the route and one edge leaving every i. Constraints (9) corresponds to the well-known sub-tours elimination constraints of the VRP (Toth and Vigo, 2002), adding the time parameter of the problem. Finally, constraints (10), (11), (12) and (13) are the domain constraints.

4. Solution Approach

Due to the NP-hard nature of the IRP, a hybrid evolutionary optimization algorithm based on two well-known meta-heuristics (Genetic Algorithm, Simulated Annealing Algorithm) is proposed to handle the problem. Since the IRP can be described as the combination of the Inventory Control and the Vehicle Routing Problems, the meta-heuristics are used as follows: The Genetic Algorithm is related to the planning phase of the IRP (inventory control problem) determining delivery times and quantities, while the Simulated Annealing Algorithm is associated with the routing phase of the IRP (vehicle routing problem) determining routes. Both Algorithm are dealt with in an iterative way to define the re-optimization phase. Hence, a repetitive procedure is applied to obtain a near-optimal feasible solution.

4.1 Planning Phase – A Genetic Algorithm Approach

Genetic Algorithm (GAs) have been developed by John Holland and his collaborators at the University of Michigan in the 1970s (Holland, 1975). They are based on the principles of biological evolution and the natural selection process of the survival of the fittest. This process actually reflects an optimization process based on an initial, randomly generated, population of solutions (population-based meta-heuristic). A solution is referred to as an *individual*, while its data structure representation corresponds to the *chromosome* or *genotype*. A chromosome consists of *genes* that represent the decision variables within a solution. One iteration of creating a new population through the optimization algorithm is called a *generation*. The population is maintained and evolved from generation to generation using genetic operators such as *evaluation, reproduction (selection), recombination (crossover)* and *mutation*. The *fitness* of each individual is associated with the evaluation function or the objective function, while the *phenotype* represents how an individual operates during the fitness assessment.

Furthermore, a selection process allows *parent* solutions with high fitness to be selected from the current population. Then, crossover and mutation operators are applied to generate children (*offspring*). In particular, the crossover operator intends to inherit some characteristics (genes) of the two parents to generate the offspring, while the mutation operator represents a slight change to a single individual. The offspring compete with the parents for their place in the next generation (survival of the fittest), thus constructing the next population. In the following subsections, a detailed description of the developed genetic approach regarding the IRP is given.

4.1.1 Illustrative Example

An illustrative example is used to demonstrate key modeling features of the Genetic Algorithm Approach. To begin with, a small sample problem of a distribution system that comprises a single supplier and six customers can be considered to illustrate the proposed chromosome representation (Figure 2).

The planning horizon is equal to six days. At the beginning of the planning horizon, all customers have zero inventory levels, whereas each customer has a daily demand. Stock-outs are not allowed, while inventory holding costs exist only at the demand points. Each customer has a sufficient maximum inventory level to satisfy his storage needs during the planning horizon. Furthermore, it is assumed that the supplier has a sufficient supply of products that can cover all customers' demands throughout the planning horizon. Table 1 provides information about the maximum inventory level as well as the inventory holding cost of each customer.

Table 1 *Inventory Information for the Illustrative Example*

Customer	Inventory Holding Cost (per unit per period)	Maximum Inventory Level
1	0.4649	115
2	0.3723	190
3	0.3545	95
4	0.4054	135
5	0.4908	60
6	0.1219	175

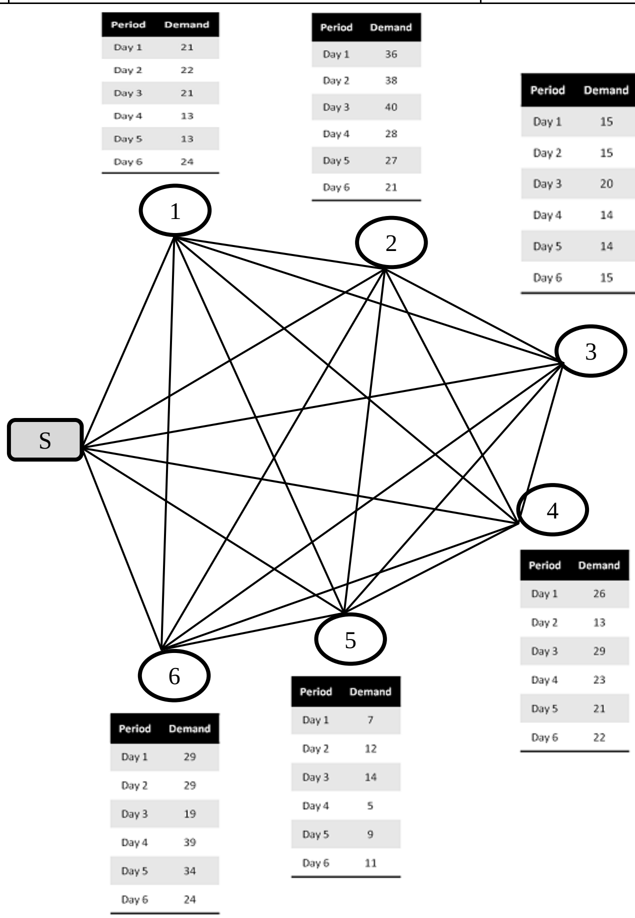


Figure 2 *Illustrative Example*

A chromosome can be represented by a two-dimensional matrix with six rows and six columns (Table 2).

The rows and the columns of the matrix correspond to the customers and the time periods of the planning horizon, respectively. Each cell of the matrix represents the total amount of product that should be delivered to a specific customer in a specific time period. For example, the total amount of product that should be delivered to customer 2 in day 3 is equal to 40. Since stock-outs are not allowed, it should be observed that each delivery quantity satisfies the current demand of the customer. If a delivery to a customer does not take place in a specific time period, the period's demand is satisfied through the available inventory from a previous delivery. For instance, the delivery quantities of period 1 for Customer 2 are enough to satisfy the demands of Period 1 and 2, respectively (43 = 21 + 22). Therefore, for each customer (row of a matrix), the sum of delivery quantities is equal to the sum of customer demand during the planning horizon.

Table 2 *Chromosome Representation (Delivery Quantities Matrix)*

	Period 1	**Period 2**	**Period 3**	**Period 4**	**Period 5**	**Period 6**
Customer 1	43	0	21	13	13	24
Customer 2	74	0	40	55	0	21
Customer 3	15	15	34	0	14	15
Customer 4	26	42	0	23	21	22
Customer 5	7	12	14	5	9	11
Customer 6	29	87	0	0	34	24

Based on a pre-defined population size, a random procedure is followed to generate the initial population. To begin with, each individual in the population is represented by a randomly generated binary matrix (Table 3). Each cell contains a 1/0 value indicating whether a customer is visited in a specific time period.

Table 3 *Binary Matrix Representation*

	Period 1	**Period 2**	**Period 3**	**Period 4**	**Period 5**	**Period 6**
Customer 1	1	0	1	1	1	1
Customer 2	1	0	1	1	0	1
Customer 3	1	1	1	0	1	1
Customer 4	1	1	0	1	1	1
Customer 5	1	1	1	1	1	1
Customer 6	1	1	0	0	1	1

Since at the beginning of the planning horizon all customers have zero initial inventory levels and stock-outs are not allowed, the first column of the binary matrix contains only 1-values. The remaining columns of the binary matrix are randomly generated. Below, an algorithm is presented that generates a binary matrix.

Algorithm1. *Generate a Binary Matrix*
Inputs: NC (number of customers), NP (number of periods) $tempBM1 \leftarrow ones(NC, 1)$ $tempBM2 \leftarrow randi([0,1], NC, NP-1), BinaryMatrix \leftarrow [tempBM1, tempBM2]$ **Output:** $BinaryMatrix$

Analytically, $ones$ creates an NC-by-1 array of ones, while $randi$ creates an NC-by-($NP - 1$) array of 1/0 values. Afterward, the algorithm combines the two arrays into one array to create the binary matrix that corresponds to an individual of the population. Given a population size, $PopSize$, this procedure can be repeated to create the initial population (Algorithm 2).

Algorithm 2. *Generate a Population of Binary Matrices*

Inputs: $NC, NP, PopSize$
    $\boldsymbol{for}\, i = 1{:}\,PopSize\,\boldsymbol{do}$
        $\boldsymbol{Call}\, Algorithm\ 1$
    $\boldsymbol{end - for}$
Output: $PopBM$ (population of binary matrices)

According to a binary matrix, a real-value matrix that consists of delivery quantities in each time period of the planning horizon can be easily produced (Algorithm 3). This two-dimensional matrix reflects the chromosome representation shown in Table 2.

Algorithm 3. *Produce Chromosome Representations of Population's Individuals*

Inputs: DM (demand matrix), $PopBM$, $PopSize$
    $\boldsymbol{for}\, i = 1{:}\,PopSize\,\boldsymbol{do}$
        $Population\{i\} \leftarrow convertBM(DM, PopBM)$
    $\boldsymbol{end - for}$
Output: $Population$

Given the customers' demands during the planning horizon and their binary matrix representations, Algorithm 3 produces real-value matrices that reflect the initial population with respect to the assumption that stock-outs are not allowed. In particular, after each iteration, $convert\ BM$ creates a delivery quantity matrix according to the demand matrix of each customer and the relative binary matrix. As a result, after each iteration, an individual is added to the population. Moreover, based on a delivery quantity matrix, inventory levels and inventory holding costs of each customer can be easily determined, as shown in Tables 4 and 5, respectively.

Analytically, Table 4 represents the inventory level I_i^t at each customer $i \in C = \{Customer1, \ldots, Customer6\}$ at the end of period $t \in T = \{Period1,\ \ldots, Period6\}$. For example, according to Figure 2, Customer6' demands are: $29\ (t = 1), 29\ (t = 2), 19\ (t = 3), 39 (t = 4), 34 (t = 5), 24 (t = 6)$. Based on the sample individual (chromosome) of Table 2, Customer6 should be visited four times during the planning horizon by delivering the following quantities: $29\ (t = 1), 87\ (t = 2), 34\ (t = 5), 24\ (t = 6)$. As a result, the inventory levels at Customer6 at the end of each time period of the planning horizon are: $I_6^1 = 0 + 29 - 29 = 0; I_6^2 = 0 + 87 - 29 = 58;\ I_6^3 = 58 + 0 - 19 = 39; I_6^4 = 39 + 0 - 39 = 0; I_6^5 = 0 + 34 - 34 = 0; I_6^6 = 0 + 24 - 24 = 0$. Therefore, for Customer6 there are only two inventory levels ($I_6^2 = 58, I_6^3 = 39$). According to the objective function (1) the inventory costs related to these time periods are: $0.1219 \times 58 = 7.0702$ and $0.1219 \times 39 = 4.7541$, respectively. In addition, the total inventory cost associated with Customer6 is $7.0702 + 4.7541 = 11.8243$ (Table 5).

Table 4 *Inventory Level Matrix*

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6
Customer 1	22	0	0	0	0	0
Customer 2	38	0	0	27	0	0
Customer 3	0	0	14	0	0	0
Customer 4	0	29	0	0	0	0
Customer 5	0	0	0	0	0	0
Customer 6	0	58	39	0	0	0

Table 5 *Inventory Holding Cost Matrix*

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6
Customer 1	10.2278	0	0	0	0	0
Customer 2	14.1474	0	0	10.0521	0	0
Customer 3	0	0	4.963	0	0	0
Customer 4	0	11.7566	0	0	0	0
Customer 5	0	0	0	0	0	0
Customer 6	0	7.0702	4.7541	0	0	0

Table 6 *Single-point Crossover Operator*

	P1	P2	P3	P4	P5	P6
C1	64	0	0	13	37	0
C2	74	0	40	55	0	21
C3	15	15	20	14	29	0
C4	26	42	0	23	21	22
C5	7	12	14	5	9	11
C6	29	87	0	0	34	24

Parent1

	P1	P2	P3	P4	P5	P6
C1	21	22	21	26	0	24
C2	114	0	0	28	27	21
C3	30	0	34	0	25	4
C4	68	0	0	44	0	22
C5	19	0	19	0	9	11
C6	29	48	0	39	34	24

Parent2

	P1	P2	P3	P4	P5	P6
C1	64	0	0	13	37	0
C2	74	0	40	55	0	21
C3	15	15	20	14	29	0
C4	68	0	0	44	0	22
C5	19	0	19	0	9	11
C6	29	48	0	39	34	24

Offspring1

	P1	P2	P3	P4	P5	P6
C1	21	22	21	26	0	24
C2	114	0	0	28	27	21
C3	30	0	34	0	25	4
C4	26	42	0	23	21	22
C5	7	12	14	5	9	11
C6	29	87	0	0	34	24

Offspring2

An important issue is the choice of an appropriate fitness function that determines the selection criterion in the IRP. The fitness quantifies the optimality of a solution (i.e., a chromosome) in the proposed hybrid evolutionary algorithm so that a particular chromosome may be ranked against all the other chromosomes. Therefore, optimal chromosomes are allowed to breed and mix their genes by any of several techniques, producing a new generation that will be even better. For the IRP, it is assumed that candidate solutions with lower total costs (inventory holding costs plus transportation costs) imply better solutions. Since the IRP is a minimization problem, the fitness for each chromosome is defined as follows: $fitness = \frac{1}{\sum_{k\in K}\sum_{t\in T}\sum_{i\in V}\sum_{j\in V,j>i} c_{ij}x_{ijk}^{t} + \sum_{t\in T}\sum_{i\in C} h_i I_i^t}$. Therefore, each individual has a probability of being selected that is proportional to its fitness. The higher the individual's fitness is, the more likely it is to be selected. In this context, the roulette-wheel selection approach is adopted as the selection process.

Algorithm 4. *Roulette-Wheel Selection*
Inputs: $fitness(1), fitness(2), \ldots, fitness(PopSize)$
$f_{sum} \leftarrow \sum_{i=1}^{i=PopSize} fitness(i)$
$Generate a uniformly distributed random number rN \in [0, f_{sum}]$
$FV \leftarrow fitness(1), iter \leftarrow 1$
$\boldsymbol{while} FV < rN \boldsymbol{do}$
$iter \leftarrow iter + 1, FV \leftarrow FV + fitness(iter)$
$\boldsymbol{end - while}$
$Parent \leftarrow iter$
Output: $Parent$

Algorithm 4 shows how to select a parent from a population of $PopSize$ individuals. To keep the population size constant across generations, suitable pairs of mates are picked. The goal is to select every time two parents to produce two offspring. This process is repeated until the population of offspring is the same as the population of parents.

Since two parents are selected, a crossover operator can be applied. For the reported chromosome representation, a single-point crossover operator as well as a double-point crossover operator has been designed and can be used randomly to produce two offspring. The two-dimensional matrix structure can be broken horizontally considering that delivery quantities for a selected set of customers will be exchanged between two parent solutions. Hence, the crossover point is relevant to a specific row of the two-dimensional matrix.

Based on the two-dimensional matrix structure, the single-point crossover indicates that one crossover position (i.e., a specific row of the matrix) is selected uniformly at random and the rows are exchanged between the individuals about this point. Then, two new offspring are produced. Consider the example of Table 6. Parent 1 and Parent 2 represent two individuals. Each individual is a delivery quantities matrix. The third row of the two-dimensional matrix structure is considered as the crossover point. The crossover operator is based on exchanging the delivery schedules between the two parents. Namely, all data beyond that crossover point (i.e., rows 4, 5 and 6 which correspond to customer 4, 5 and 6, respectively) in either parent is swapped between the two parents. As a result, Parent 1 and Parent 2 exchange rows 4, 5 and 6 with each other, thus producing two offspring. Offspring1 contains the delivery schedules related to customers 1, 2 and 3 from Parent1 and the delivery schedules related to customer 4, 5 and 6 from Parent2. Similarly, Offspring2 contains the delivery schedules related to customers 1, 2 and 3 from Parent2 and the delivery schedules related to customers 4, 5 and 6 from Parent1. It is worth noting that the delivery schedules for each customer defined in parents will remain unchanged in the resultant offspring (Offspring1 and Offspring2).

The algorithm that shows the functionality of the single-point crossover operator is presented below (Algorithm 5). Assuming that two parents are selected, $parent(1)$ and $parent(2)$ from a given population, $population$, a crossover point is randomly generated from $[2, NC - 1]$, where NC is the given number of customers. Since the crossover point is known, two offspring are produced. The first offspring, O_1, as well as the second one, O_2, maintain the first C$rossPoint$ rows of P_1 and P_2, respectively. In addition, the remaining rows of P_1, $CrossPoint + 1, \dots, NC$, are copied to O_2, while the remaining rows of P_2 are copied to O_1. Furthermore, to guarantee the continuity of the process, the relative binary matrices for O_1 and O_2 are produced, called O_1^BM andO_2^{BM}, respectively.

Algorithm 5. *Single-Point Crossover*

Inputs: $PopBM, Population, Parent(1), Parent(2), NC$

$P_1 \leftarrow Parent(1), P_2 \leftarrow Parent(2)$

$CrossPoint \leftarrow randomintegerfrom[2, NC - 1]$

$O_1 \leftarrow Population\{P_1\}, O_2 \leftarrow Population\{P_2\}, O_1^{BM} \leftarrow PopBM\{P_1\}$

$O_2^{BM} \leftarrow PopBM\{P_2\}$

$O_1(CrossPoint + 1: end, :) \leftarrow Population\{P_2\}(CrossPoint + 1: end, :)$

$O_2(CrossPoint + 1: end, :) \leftarrow Population\{P_1\}(CrossPoint + 1: end, :)$

$O_1^{BM}(CrossPoint + 1: end, :) \leftarrow PopBM\{P_2\}(CrossPoint + 1: end, :)$

$O_2^{BM}(CrossPoint + 1: end, :) \leftarrow PopBM\{P_1\}(CrossPoint + 1: end, :)$

Outputs: $O_1, O_2, O_1^{BM}, O_2^{BM}$

In the double-point crossover operator, two crossover positions are selected uniformly at random and the rows are exchanged between the individuals between these points. Then, two new offspring are produced. Consider the following example.

Table 7 *Double-point Crossover Operator*

	P1	P2	P3	P4	P5	P6
C1	64	0	0	13	37	0
C2	74	0	40	55	0	21
C3	15	15	20	14	29	0
C4	26	42	0	23	21	22
C5	7	12	14	5	9	11
C6	29	87	0	0	34	24

Parent1

	P1	P2	P3	P4	P5	P6
C1	21	22	21	26	0	24
C2	114	0	0	28	27	21
C3	30	0	34	0	25	4
C4	68	0	0	44	0	22
C5	19	0	19	0	9	11
C6	29	48	0	39	34	24

Parent2

	P1	P2	P3	P4	P5	P6
C1	64	0	0	13	37	0
C2	74	0	40	55	0	21
C3	30	0	34	0	25	4
C4	68	0	0	44	0	22
C5	7	12	14	5	9	11
C6	29	87	0	0	34	24

Offspring1

	P1	P2	P3	P4	P5	P6
C1	21	22	21	26	0	24
C2	114	0	0	28	27	21
C3	15	15	20	14	29	0
C4	26	42	0	23	21	22
C5	19	0	19	0	9	11
C6	29	48	0	39	34	24

Offspring2

This example depicts two crossover positions, row 2 and row 5. Therefore, rows 3 and 4 of the Parent 1, P_1, are copied to the Offspring 2, O_2, while rows 3 and 4 of the Parent 2, P_2, are copied to the Offspring 1, O_1. Analogous to the simple-point crossover operator, an algorithm of the double-point crossover operator is presented below (Algorithm 6).

Algorithm 6. *Double-Point Crossover*
Inputs: $PopBM, Population, Parent(1), Parent(2), NC$
$P_1 \leftarrow Parent(1), P_2 \leftarrow Parent(2)$
$CrossPoint_1 \leftarrow randomintegerfrom[2, NC-1]$
$CrossPoint_2 \leftarrow randomintegerfrom[2, NC-1]$
$sort(CrossPoint_1, CrossPoint_2)$ $//CrossPoint_1 \neq CrossPoint_2$
$a \leftarrow CrossPoint_1$, b$\leftarrow$ $CrossPoint_2$
$O_1 \leftarrow Population\{P_1\}, O_2 \leftarrow Population\{P_2\}, O_1^{BM} \leftarrow PopBM\{P_1\}$
$O_2^{BM} \leftarrow PopBM\{P_2\}$
$O_1(a+1{:}b-1,:) \leftarrow Population\{P_2\}(a+1{:}b-1,:)$
$O_2(a+1{:}b-1,:) \leftarrow Population\{P_1\}(a+1{:}b-1,:)$
$O_1^{BM}(a+1{:}b-1,:) \leftarrow PopBM\{P_2\}(a+1{:}b-1,:)$
$O_2^{BM}(a+1{:}b-1,:) \leftarrow PopBM\{P_1\}(a+1{:}b-1,:)$
Outputs: $O_1, O_2, O_1^{BM}, O_2^{BM}$

After the crossover, an individual is subjected to mutation. In particular, the mutation prevents the algorithm from being trapped in a local minimum. Therefore, through the crossover, a current solution is exploited to find better ones, whereas the mutation is supposed to help to explore the whole search space. In the context of the proposed solution approach, the mutation operator presented by Abdelmaguid and Dessouky (2006), called the backward delivery exchange, is adopted. The backward delivery exchange process is chosen due to the restriction that stock-outs are not allowed. Accordingly, part of a customer's delivery amount can be transferred only to a preceding period. Table 8 illustrates an example of using the backward delivery exchange operator.

Table 8 *Backward Delivery Exchange Operator*

	Period 1	**Period 2**	**Period 3**	**Period 4**	**Period 5**	**Period 6**
Customer 1	43	0	21	13	13	24
Customer 2	74	0	40	55	0	21
Customer 3	15	15	34	0	14	15
Customer 4	26	42	0	23	21	22
Customer 5	7	~~12~~ 22	~~14~~ 4	5	9	11
Customer 6	29	87	0	0	34	24

The example shows that the delivery quantity for Customer 5 scheduled in period 3 is reduced by 10 units, and this amount is transferred to period 2.

Algorithm 7. *Backward Delivery Exchange Operator*

Input: O (child solution)
$\boldsymbol{if}\big(rand \leftarrow U(0,1)\big) < MutationRate\,\boldsymbol{then}, RandNC \leftarrow randomintegerfrom[1, NC]$
$\boldsymbol{for}\,mutInd = 1{:}\,RandNC\,\boldsymbol{do}$
$RandPer \leftarrow randomintegerfrom[2, NP]$
$\boldsymbol{while}\,sum(O(:, RandPer)) = 0\ \boldsymbol{do}$
$RandPer \leftarrow randomintegerfrom[2, NP]$
$\boldsymbol{end - while}$
$ListCust \leftarrow find\big(O(:, RandPer)\big), n \leftarrow length(ListCust)$
$RandMutCust \leftarrow randomintegerfrom[1, n]$
$DelX \leftarrow O(ListCust(RandMutCust), RandPer)$
$BcDelX \leftarrow randomintegerfrom[1, DelX]$
$PrecPer \leftarrow findPrecedingPeriod(RandPer)$
$tempDel \leftarrow O(ListCust(RandMutCust), PrecPer)$
$O(ListCust(RandMutCust), PrecPer) \leftarrow tempDel + BcDelX$
$O(ListCust(RandMutCust), RandPer) \leftarrow DelX - BcDelX$
$\boldsymbol{end - for}$
$\boldsymbol{end - if}$
Output: O (child solution after mutation)

Below, the algorithm of the backward delivery exchange operator is presented (Algorithm 7). An important parameter in the mutation process is the mutation probability, $MutationRate$, which decides how often parts of a chromosome will be mutated. Since a mutation takes place, a random integer, $RandNC$, is generated in the interval from$[1, NC]$, where NC indicates the number of the customers. Then, for $RandNC$ times, the following process is repeated. A period, $RandPer$, is selected randomly. If no deliveries are scheduled for this period, $RandPer$ is re-generated randomly. Afterwards, $find$ returns the customers, $ListCust$, that are scheduled to be visited in the $RandPer$ time period of the planning horizon. A customer, $RandMutCust$, from $ListCust$ is randomly selected and his scheduled delivery amount, $DelX$, is saved. Next, the amount that could be transferred to a preceding period, $BcDelX$, is randomly selected in the interval from $[1, DelX]$. From previous periods where a customer has scheduled deliveries, the nearest period, $PrecPer$, is selected to transfer $BcDelx$ units of product. Subsequently, the scheduled delivery quantity in period $RandPer$ is reduced by $BcDelX$ units, and this amount is transferred to period $PrecPer$.

4.2 Routing Phase – A Simulated Annealing Algorithm Approach

Since the Genetic Algorithm focuses only on the planning phase by determining the delivery times and quantities, the vehicle routes should be constructed. The routing phase is related to the usage of a Simulated Annealing Algorithm for solving a vehicle routing problem for each time period of the planning horizon where delivery quantities have been scheduled. The Simulated Annealing Algorithm is a nature-inspired optimization algorithm introduced by Kirkpatrick et al. (1983). Contrary to Genetic Algorithm, it is a single-individual stochastic algorithm, as it does not involve a population of candidate solutions. The algorithm mimics the annealing process of heating and cooling a material in order to re-crystallize it (Talbi, 2009). In particular, the annealing process starts with an initial system state at a very high temperature, which is slowly decreased to obtain a strong crystalline structure. The strength of the structure depends on the rate of decrease, which is subjected to a cooling process until it converges to an equilibrium state (steady frozen state). However, to reach an equilibrium state at each temperature, a number of sufficient transitions must be applied.

Similarly, the Simulated Annealing algorithm (Algorithm 8) consists of two cycles, the external and the internal cycle. The algorithm begins with an initial feasible solution, x_0, and a high temperature T_{max} and proceeds in $EXTcyc$iterations (external cycle). Then, the algorithm proceeds in $INTcyc$ iterations (internal cycle). Throughout the internal cycle, the temperature is constantly trying to converge to an equilibrium state at the end of $INTcyc$ iterations. At each iteration of the internal cycle, a neighboring solution, x, is generated by perturbing the current solution. A cost function, $CostFunction()$, exists to measure the quality of each solution. If the cost of the neighboring solution, $CostFunction(x)$, is less than the cost of the current solution, $CostFunction(x_0)$, it is accepted. Otherwise, it is accepted with probability$e^{-\frac{\Delta E}{T}}$, where T is a control parameter (temperature) and ΔE represents the difference in the objective value between the current solution and the generated neighboring solution. The control parameter T is decreased gradually through the external cycle. The temperature is updated using a geometric schedule that corresponds to the formula $T \leftarrow alpha \times T$, where $alpha \in [0,1]$. Therefore, as the algorithm progresses, the probability that a non-improving generated neighboring solution is accepted decreases. The set of parameters related to the high value of control parameter (*temperature*), T_{max}, the rate of decrease (*cooling rate*), $alpha$, and the stopping condition of the internal ($INTcyc$ iterations) as well as external cycle ($EXTcyc$ iterations) of the algorithm is called the *annealing (cooling) schedule*.

In terms of the optimization process, the annealing schedule controls the transition from the exploration to the exploitation. Particularly, at the beginning of the algorithm, the temperature has a high value, which is decreased until a final temperature is reached. This final temperature is typically close to zero. As a consequence, at the beginning of the algorithm, the exploration is high and the exploitation is low, while at the end of the algorithm, the exploitation is high and the exploration is low. The main objective is to obtain a balance between exploration and exploitation to sufficiently explore the search space and simultaneously exploit good solutions.

Assume an individual in a population obtained by the planning phase described in section 4.1. The Simulated Annealing algorithm should be applied to each time period of the planning horizon where scheduled delivery quantities exist (Table 9).

Path representation is the most natural way of representing the routes of a VRP. Since a VRP consists of one or more routes, the length of each path is variable. On account of this, a dynamic variable, x, can be used to represent the solution of the VRP. x Contains all the routes of a specific time period of the planning horizon. For instance, in the first time period $x = \{x.R_1, x.R_2, x.R_3\}$, where (a) $x.R_1 = [0,2,13,,8,3,10,0]$ is the first route, (b) $x.R_2 = [0,4,11,15,12,1,0]$ is the second route and (c) $x.R_3 = [0,6,7,9,14,5,0]$ is the third route. The zero value in each row vector represents the supplier, while the other numbers represent the customers.

Algorithm 8. *Simulated Annealing Algorithm*

Inputs: $T_{max}, EXTcyc, INTcyc, alpha$

$x_0 \leftarrow GenerateInitialSolution(\ \ )\ //\ Algorithm\ 9$

$T \leftarrow T_{max}, i \leftarrow 1$

$\boldsymbol{while}\ i \leq EXTcyc\ \boldsymbol{do}$

$\quad j \leftarrow 1$

$\quad \boldsymbol{while}\ j \leq INTcyc\ \boldsymbol{do}$

$\quad\quad x \leftarrow CreateNeighboringSolution(x_0)//\ Algorithm\ 10$

$\quad\quad \boldsymbol{if}\ CostFunction(x) < CostFunction(x_0)\ \boldsymbol{then}$

$\quad\quad\quad x_0 \leftarrow x$

$\quad\quad \boldsymbol{else}$

$\quad\quad \Delta E \leftarrow CostFunction(x_0) - CostFunction(x)$

$\quad\quad\quad randN \leftarrow U(0,1)$

$\quad\quad \boldsymbol{if}\ randN < e^{-\frac{\Delta E}{T}}\ \boldsymbol{then}, x_0 \leftarrow x, \boldsymbol{end-if}$

$\quad\quad \boldsymbol{end-if}$

$\quad\quad j \leftarrow j + 1$

$\quad \boldsymbol{end-while}$

$\quad T \leftarrow alpha * T, i \leftarrow i + 1$

$\boldsymbol{end-while}$

Output: *best solution found*

Table 9 *Solving a VRP Problem at Each Time Period of the Planning Horizon*

	Period 1	**Period 2**	**Period 3**	**Period 4**	**Period 5**	**Period 6**
Customer 1	64	0	0	26	0	24
Customer 2	36	38	68	0	44	4
Customer 3	30	0	20	43	0	0
Customer 4	39	0	52	0	43	0
Customer 5	19	0	39	0	0	0
Customer 6	58	0	116	0	0	0
Customer 7	23	0	17	8	23	0
Customer 8	110	0	0	25	0	0
Customer 9	30	0	0	0	17	0
Customer 10	10	13	0	13	4	0
Customer 11	19	21	0	0	0	0
Customer 12	34	0	52	0	0	0
Customer 13	14	0	14	0	16	0
Customer 14	51	0	0	69	0	0
Customer 15	38	0	0	7	18	0
	VRP	VRP	VRP	VRP	VRP	VRP

In order to generate an initial solution to start solving a VRP with Simulated Annealing algorithm, a random approach is followed, $GenerateInitialSolution()$. The approach iterates over a pre-defined list of customers that will be visited in a specific time period according to the planning phase. The algorithm (Algorithm 9) proceeds as follows. If there are n customers in the pre-defined list, a customer is selected randomly to start creating a route. Each route corresponds to a specific vehicle of a fleet with capacity Q. Moreover, each customer who is added in a route should have delivery quantity, $DQ_{SelectedCustomer}$, such that it does not exceed the vehicle's capacity, VC, whereas this customer is excluded from the list since he is associated with a specific route. If a customer's delivery quantity is greater than the remaining vehicle's capacity, VC, a new route is designed that is related to a new vehicle with capacity $VC = Q$. Finally, the combination of the routes construct the random initial non-optimal feasible solution, x_0, of the VRP.

Algorithm 9. *Generate an Initial VRP Solution*

Inputs:$ListOfCustomers, Q$
$n \leftarrow length(ListOfCustomers), i \leftarrow 1, j \leftarrow 1, VC \leftarrow Q$
$\boldsymbol{while} i \leq n \boldsymbol{do}$
$SelectedCustomer \leftarrow generaterandomintegerfrom\ [ListOfCustomers]$
$\boldsymbol{if} DQ_{SelectedCustomer} \leq VC\boldsymbol{then}$
$\boldsymbol{if} DQ_{SelectedCustomer} = VC\boldsymbol{then}$
$R_j \leftarrow addSelectedCustomer, UpdateListOfCustomers,$
$x_0 \leftarrow R_j, j \leftarrow j+1, VC \leftarrow Q, i \leftarrow i+1$ $\boldsymbol{else}$
$R_j \leftarrow addSelectedCustomer, UpdateListOfCustomers$
$VC \leftarrow VC - DQ_{SelectedCustomer}, i \leftarrow i+1$
$\boldsymbol{end-if}$
$\boldsymbol{else}$
$x_0 \leftarrow R_j, j \leftarrow j+1, VC \leftarrow Q$
$R_j \leftarrow addSelectedCustomer, UpdateListOfCustomers$
$VC \leftarrow VC - DQ_{SelectedCustomer}, i \leftarrow i+1$
$\boldsymbol{end-if}$
$\boldsymbol{end-while}$
$updatex_0$
Output:$x_0 = \{x_0.R_1, x_0.R_2, \ldots\}$

The objective of the Simulated Annealing algorithm is to minimize the cost associated with all proposed routes of a specific time period of the planning horizon. Therefore, if a solution x consists of k routes, the cost function, $CostFunction$, is equal to: $\sum_{j=1}^{j=k} Cost(x.R_j)$.

At each iteration of the internal cycle of the Simulated Annealing algorithm, a neighboring solution,x, is generated by perturbing the current solution, $CreateNeighboringSolution(x_0)$ (Algorithm 10). The generation of the neighboring solution is based on a random selection among three inter-route improvement Algorithm: (a) the move improvement algorithm, (b) the swap improvement algorithm and (c) the cyclic improvement algorithm. The three Algorithm attempt to reduce the total route length by moving one or more customers to a different route. It is worth noting that a move is feasible if the demand of the moved customer does not violate the vehicle capacity on the route it is moved to. All of the Algorithm are analytically described by Goetschalckx (2011).

Algorithm 10. *Generate a Neighboring Solution*
Input: x_0
$randN \leftarrow generaterandomintegerfrom$ [1,3]
$\boldsymbol{if} randN = 1$ $\boldsymbol{then}, x \leftarrow move(x_0), \boldsymbol{end-if}$
$\boldsymbol{if} randN = 2$ $\boldsymbol{then}, x \leftarrow swap(x_0), \boldsymbol{end-if}$
$\boldsymbol{if} randN = 3$ $\boldsymbol{then}, x \leftarrow cyclic(x_0), \boldsymbol{end-if}$
Output: $x = \{x.R_1, x.R_2, \dots\}$

Re-optimization Phase – A Hybrid Approach

Both approaches that are presented in sections 4.1 (Genetic Algorithm) and 4.2 (Simulated Annealing) are dealt with in an iterative way, thus constructing a hybrid evolutionary optimization algorithm (Algorithm 11) that is related to a re-optimization phase. Hence, a repetitive procedure is applied to obtain a near-optimal feasible solution. The algorithm starts by creating the IRP model based on a specific IRP data set, $IRPInstance$. Then, Algorithm 2 is called to generate an initial population of $PopSize$ individuals as far as the random binary matrices are concerned. Algorithm 3 is called to generate the population of the genetic algorithm based on the random binary matrices. Since an initial population has been constructed, a population of VRP problems is created. Each element of the population consists of a set of VRPs that correspond to each time period of the planning horizon of each individual of the genetic algorithm population. Consequently, for each individual of the population, Algorithm 8 is used to solve a VRP problem for each time period of the planning horizon where scheduled delivery quantities exist. Since the delivery quantities and times as well as the VRP solutions are available for each individual of the population, respective populations containing information about the inventory levels, $PopIM$, the inventory costs, $PopICM$, and the vehicle routing costs, $PopVRPCM$, can be created. In addition, the total inventory routing cost is calculated for each individual of the population since $PopICM$ and $PopVRPCM$ are available. With respect to the minimum inventory routing cost, the best individual of the population is selected, $BestIRPSol$, whereas the population is sorted.

After initialization, the algorithm proceeds as follows. For each generation, an internal cycle takes place to produce the offspring. At each iteration of the internal cycle, two parents are selected according to their fitness using Algorithm 4. Algorithm 5 or Algorithm 6 is used randomly to apply a crossover operator to produce two offspring. For each offspring, a mutation operator may be applied using Algorithm 7. After the internal cycle, a new population has been created consisting of both parents, $Population$, and offspring, $newPopulation$. Furthermore, to avoid duplicate individuals, a procedure called $replaceDuplicatesIRP()$ is applied. This procedure uses Algorithm 7, applying the proposed mutation operator to duplicate individuals. Afterward, the new best IRP solution, $BestIRPSol2$, is calculated and compared with the previous best IRP solution, $BestIRPSol$. If the second solution is better, it is accepted. Otherwise, the new population is sorted and only the first $PopSize$ individuals are selected to keep the population size constant from one generation to the next.

Algorithm 11. *Hybrid Evolutionary Optimization Algorithm*

```
Inputs: PopSize, MaxGen, MutationRate, Q, IRPInstance
IRPmodel ← LoadIRPData(IRPInstance), PopBM ← call Algorithm 2
Pop ← call Algorithm 3
pVRPs ← createVRPs(IRPmodel, Pop, Q)
for each element in pVRPs do, PopVRPs ← call Algorithm 8, end − for
PopIM ← createPopIM(IPRmodel, Pop), PopICM ← createPopICM(IRPmodel, PopIM)
PopVRPCM ← createPopVRPCM(PopVRPs)
PopIRPCM ← createPopIRPCM(PopICM, PopVRPCM)
BestIRPSol ← [PopBM{best}, Pop{best}, PopVRPs{best}, PopIRPCM{best}]
[PopulationBM, Population, PopulationIRPCM] ← popSort(PopBM, Pop, PopIRPCM)
selectedParents ← zeros(1,2)
for indx = 1: MaxGen do
        for k = 1: 2: PopSize do
        Fitness ← CreateFitness(PopulationIRPCM)
                        selectedParents(1) ← call Algorithm 4
                selectedParents(2) ← call Algorithm 4, randN ← U(0,1)
                if randN < 0.5 then
                        call Algorithm 5
                else
                        call Algorithm 6
                end − if
                for i = k: k + 1 do, call Algorithm 7, end − for
        end − for
        PopA ← [PopulationBM, newPopulationBM]
                        PopB ← [Population, newPopulation]
        [PopA, PopB] ← replaceDuplicatesIRP(PopA, PopB), PopBM ← PopA
                                Pop ← PopB
        pVRPs ← createVRPs(IRPmodel, Pop, Q)
        for each element in pVRPs do, PopVRPs ← call Algorithm 8, end − for
PopIM ← createPopIM(IPRmodel, Pop)
PopICM ← createPopICM(IRPmodel, PopIM)
PopVRPCM ← createPopVRPCM(PopVRPs)
                PopIRPCM ← createPopIRPCM(PopICM, PopVRPCM)
    BestIRPSol2 ← [PopBM{best}, Pop{best}, PopVRPs{best}, PopIRPCM{best}]
        if BestIRPSol2 < BestIRPSol then
                BestIRPSol ← BestIRPSol2
else
        [PopulationBM, Population, PopulationIRPCM] ←
popSort(PopBM, Pop, PopIRPCM)
        PopulationBM ← PopulationBM(1: PopSize)
                        Population ← Population(1: PopSize)
        end − if
end − for
Output: BestIRPSol
```

Based on the example presented in Section 4.2., Table 10 presents the routes that take place in each time period of the planning horizon. If the number of generations is increased (e.g., 100 instead of 50), then a better solution is obtained (Table 11).

Table 10 *Cost Information and Routes for the Sample Problem – 50 Generations*

Routes of Period 1	Routes of Period 2	Routes of Period 3
Route 1: 0-2-13-8-3-10-0 Route 2: 0-4-11-15-12-1-0 Route 3: 0-6-7-9-14-5-0	Route 1: 0-2-10-11-0	Route 1: 0-2-3-12-4-0 Route 2: 0-6-13-7-5-0
Routes of Period 4	**Routes of Period 5**	**Routes of Period 6**
Route 1: 0-1-10-15-3-8-7-14-0	Route 1: 0-7-9-13-2-10-15-4-0	Route 1: 0-2-1-0
Total VRP Cost	**Total Inventory Control Cost**	**Total IRP Cost**
736.7608	245.1916	981.9524

Table 11 *Cost Information and Routes for the Sample Problem – 100 Generations*

Routes of Period 1	Routes of Period 2	Routes of Period 3
Route 1: 0-5-14-9-13-2-6-0 Route 2: 0-1-3-8-7-0 Route 3: 0-4-11-15-12-10-0	Route 1: 0-6-7-2-0 Route 2: 0-4-15-0	Route 1: 0-2-8-13-9-14-5-0
Routes of Period 4	**Routes of Period 5**	**Routes of Period 6**
Route 1: 0-1-10-3-2-7-5-6-0	Route 1: 0-2-7-9-14-5-0	Route 1: 0-6-8-3-12-1-0
Total VRP Cost	**Total Inventory Control Cost**	**Total IRP Cost**
711.8943	231.7815	943.6758

5. Computational Experiments and Results

This section presents the computational results of the proposed hybrid evolutionary optimization algorithm described in Section 4. The algorithm was developed in the MATLAB programming language and executed on a DELL personal computer with an Intel® Core™ i3-2120, clocked at 3.30 GHz, a microprocessor with 4 GB of RAM memory under the operating system Microsoft Windows 7 Professional. As mentioned in Section 2, new benchmark instances were designed. Consequently, the efficiency and the effectiveness of the proposed algorithm cannot be compared to other published IRP studies using benchmark instances previously introduced. This is due to the differentiated manner in which the proposed algorithm operates based on the assumptions presented in Section 2 and 3, respectively. However, this section validates the evolutionary algorithm and then evaluates its performance by comparing the algorithm's solutions with solutions obtained by solving a VRP problem for each time period of the planning horizon based on the known demands (the planning phase is ignored). The algorithm has been tested on a newly introduced set of 18 IRP benchmark instances described in the following. All benchmark instances and their computational results are available at http://www.msl.aueb.gr/files/GaSaIRP.zip.

5.1 Set of Benchmark Instances and Parameter Setting

New datasets have been developed by generalizing the well-known dataset P of Augerat et al. (1998). These datasets are divided into two classes. The first class (Class A) contains the instances with planning horizon $H = 6$ time periods (days) and a high inventory holding cost of the customers, $h_i \in [0.1,0.5] \forall i \in C$. The second class (Class B) contains the instances with planning horizon $H = 6$ time periods (days) and low inventory holding costs of the customers,

$h_i \in [0.01,0.05] \forall i \in C$. The datasets are named in the form of "IRP_nX_pY_HC" (first class) or "IRP_nX_pY_LC" (second class) strings, where "X" stands for the number of customers and "Y" stands for the number of time periods. For instance, the problem IRP_n15_p6_HC represents a test problem with 15 customers, a planning horizon of 6 days and high inventory holding costs at the customers. Different problem sizes, based on the total number of customers, were designed in each class. Class A contains problems with 15, 20, 22, 39, 44, 50, 54, 59, 64, 69, 75 and 100 customers, while the Class B contains problems with 15, 20, 22, 39, 44 and 50 customers. Vertex coordinates are kept the same as in the study by Augerat et al. (1998). The distance matrix is obtained by calculating the Euclidean distances (symmetric cost matrix). Demand exists for each customer at each time period of the planning horizon. Customer demand at each time period was generated according to the Poisson distribution, $Poisson(\lambda)$, where λ is the rate parameter. For each customer, the rate parameter is equal to his demand in the single-period VRP problem of Augerat et al. (1998). An unlimited fleet of identical vehicles with capacity Q is available for the distribution of the product. The vehicle capacity varies from 200 to 300 units of product. At the beginning of the planning horizon, all customers have zero inventory levels. Each customer has a sufficient maximum inventory level to satisfy his storage needs during the planning horizon. Namely, for each customer $i \in C$, $\sum_{t=1}^{t=H} d_i^t \leq U_i$. Finally, the supplier has a sufficient supply of products that can cover customers' demands throughout the planning horizon.

The proposed hybrid evolutionary algorithm has seven parameters to be set. Four of the parameters are associated with the Simulated Annealing algorithm. T_{max} Determines the initial value of the temperature. $EXTcyc$and$INTcyc$ are the maximum number of iterations for the external and internal cycle, respectively. In addition, $alpha$ reflects the cooling rate of the geometric schedule. The other three parameters are related to the Genetic Algorithm. Particularly, $PopSize$ defines the size of the population, $MaxGen$ sets the maximum number of generations (i.e., maximum number of iterations), while $MutationRate$ corresponds to the mutation rate. Based on the minimal cost criterion, the value of each parameter is determined after some experiments in the context of the VRP and the Inventory Control Problem, respectively. The values of the above parameters for each instance are presented in Table 12.

Table 12 *Parameters of the Hybrid Evolutionary Optimization Algorithm*

Instance	T_{max}	EXTcyc	INTcyc	alpha	PopSize	MaxGen	MutationRate
Class A Instances							
IRP_n15_p6_HC	100	1500	100	0.98	10 (20)	50 (100)	0.08
IRP_n20_p6_HC	100	1500	100	0.98	10	50	0.08
IRP_n22_p6_HC	100	1500	100	0.98	10	50	0.08
IRP_n39_p6_HC	100	1500	100	0.98	10	50	0.08
IRP_n44_p6_HC	100	1500	100	0.98	10	50	0.08
IRP_n50_p6_HC	100	1500	100	0.98	10	50	0.08
IRP_n54_p6_HC	100	1500	100	0.98	10 (20)	50 (70)	0.08
IRP_n59_p6_HC	100	1500	100	0.98	10	50	0.08
IRP_n64_p6_HC	100	1500	100	0.98	10	50	0.08

IRP_n69_p6_HC	100	1500	100	0.98	10	50	0.08
IRP_n75_p6_HC	100	1500	100	0.98	10 (20)	50 (70)	0.08
IRP_n100_p6_HC	100	1500	100	0.98	10 (20)	50 (70)	0.08
Class B Instances							
IRP_n15_p6_LC	100	1500	100	0.98	10	50	0.08
IRP_n20_p6_LC	100	1500	100	0.98	10	50	0.08
IRP_n22_p6_LC	100	1500	100	0.98	10	50	0.08
IRP_n39_p6_LC	100	1500	100	0.98	10	50	0.08
IRP_n44_p6_LC	100	1500	100	0.98	10	50	0.08
IRP_n50_p6_LC	100	1500	100	0.98	10	50	0.08

5.2 Results

This section presents the computational results for the 12 and 6 instances of Class A and B, respectively. Since the algorithm cannot be compared to other published IRP studies, the best solution obtained from the proposed algorithm (IRP) is compared to the best solution obtained if the planning phase is ignored ($p-VRP$). In the aftermath of ignoring the planning phase, a VRP problem needs to be solved for each day of the planning horizon according to daily demand. The proposed Simulated Annealing algorithm for the routing phase is then used to solve a daily VRP problem through the planning horizon. To compare the results, the following gap percentage formula is used: $Gap\ (\%) = (Sol_{IRP} - Sol_{p-VRP}) \times \frac{1}{Sol_{p-VRP}} \times 100$. The Sol_{p-VRP} corresponds to the solution obtained by solving the daily VRPs according to the known daily demands, while the Sol_{IRP} determines the solution obtained by applying the proposed hybrid evolutionary optimization algorithm. Since the Sol_{IRP} is compared with the Sol_{p-VRP}, a positive gap means that the Sol_{p-VRP} is outperformed. The computational results obtained are summarized in Table 13 (Class A instances) and 14 (Class B instances). For the $p-VRP$ problem, the total vehicle routing cost is presented, whereas for the IRP problem, the total cost is separated in terms of its transportation and inventory cost. In addition, the last column of the table shows the gap between the two problems reflecting the respective relative error.

Based on Table 13, it can be concluded that better solutions are obtained when the planning phase is considered. The ability of each customer to have storage enables a significant decrease in the vehicle routing cost, reducing the total number of routes during the planning horizon. For 8 of the 12 instances, the evolutionary algorithm provides better solutions with gaps in the interval of -27.4365 percent to -0.4213 percent. It should not pass unnoticed that even in cases where the $p-VRP$ provides better solutions, a change in parameters of the evolutionary algorithm, such as population size and maximum number of generations (see Table 4), results in a gap improvement. Specifically, for the first, the eleventh and the twelfth instance of Class A, the gap was improved by 24%, 37.16% and 116.04%, respectively. In particular, in the last case, the improvement of the gap was such that the best solution of the instance was improved significantly. Despite the better solution obtained from the $p-VRP$, the change in the parameters led to the evolutionary algorithm providing an even better solution.

Table 13 *Experimental Results (First Class of Instances)*

Instance	p-VRP	IRP			p-VRP – IRP
	Vehicle Routing Cost	Vehicle Routing Cost	Inventory Holding Cost	Total Cost	Gap (%)
IRP_n15_p6_HC	1141.4608	736.7608	245.1916	981.9524	-13.9741
		711.8943	231.7815	943.6758	-17.3274
IRP_n20_p6_HC	1313.139	947.7804	323.2335	1271.0139	-3.2080
IRP_n22_p6_HC	1348.1933	1073.4663	245.6849	1319.1511	-2.1542
IRP_n39_p6_HC	2481.3249	1638.3222	648.6909	2287.0132	-7.8310
IRP_n44_p6_HC	2832.1262	1311.7218	743.3693	2055.0912	-27.4365
IRP_n50_p6_HC	2934.1541	1546.7005	780.6426	2327.3431	-20.6810
IRP_n54_p6_HC	2991.2924	1931.812	964.661	2896.4729	-3.1669
		2281.1826	801.8551	3083.0377	3.0671
IRP_n59_p6_HC	3472.2657	2560.0989	958.1427	3518.2415	1.3241
IRP_n64_p6_HC	3707.7601	2691.5032	1143.1868	3834.69	3.4234
IRP_n69_p6_HC	4014.4543	3015.5997	1215.7251	4231.3248	5.4022
IRP_n75_p6_HC	3923.5927	3029.7364	1208.7383	4238.4747	8.0253
		2930.6974	1190.7582	4121.4556	5.0429
IRP_n100_p6_HC	4854.3807	3791.122	1190.755	4981.877	2.6264
		3645.1169	1188.8117	4833.9286	-0.4213

Furthermore, Table 14 shows the computational results related to the instances of Class B. As can be observed, in all cases, the evolutionary algorithm provides better solutions than the $p - VRP$, with gaps in the interval of -49.9496 percent to -24.8749 percent. The results indicate that if a small inventory holding cost is applied to each customer, better solutions can be obtained, significantly reducing the total vehicle routing cost and designating the importance of integrating supply chain activities.

Table 14 *Experimental Results (Second Class of Instances)*

Instance	p-VRP	IRP			p-VRP – IRP
	Vehicle Routing Cost	Vehicle Routing Cost	Inventory Holding Cost	Total Cost	Gap (%)
IRP_n15_p6_LC	1141.4608	664.3223	25.0305	689.3528	-39.6078
IRP_n20_p6_LC	1313.139	774.6633	36.3499	811.0132	-38.2386
IRP_n22_p6_LC	1348.1933	986.0693	26.7627	1012.832	-24.8749
IRP_n39_p6_LC	2481.3249	1406.3136	63.409	1469.7226	-40.7686
IRP_n44_p6_LC	2832.1262	1428.9866	73.6921	1502.6787	-46.9417
IRP_n50_p6_LC	2934.1541	1276.7132	191.8418	1468.555	-49.9496

To illustrate in more detail the behavior of the proposed algorithm, more information is presented about the vehicles (number of routes) used in each time period of the planning horizon in Tables 15 and 16.

Table 15 *Number of Vehicles used during the Planning Horizon (First Class of Instances)*

Instance	p-VRP							IRP						
	P1	P2	P3	P4	P5	P6	No. of Routes	P1	P2	P3	P4	P5	P6	No. of Routes
IRP_n15_p6_HC	2	2	2	2	2	2	12	3	1	2	1	1	1	9
								3	2	1	1	1	1	9
IRP_n20_p6_HC	2	2	2	2	2	2	12	4	2	1	2	1	1	11
IRP_n22_p6_HC	2	2	2	2	2	2	12	4	2	1	2	2	1	12
IRP_n39_p6_HC	3	3	3	3	3	3	18	6	2	3	2	1	1	15
IRP_n44_p6_HC	3	3	3	3	3	3	18	14	0	1	1	0	0	16
IRP_n50_p6_HC	3	3	3	3	3	3	18	16	0	2	1	1	0	20
IRP_n54_p6_HC	4	4	4	4	4	4	24	16	2	3	1	1	1	24
								7	4	4	3	4	2	24
IRP_n59_p6_HC	4	4	4	5	4	4	25	8	4	4	5	3	2	26
IRP_n64_p6_HC	5	5	5	5	5	4	29	12	5	5	3	2	1	28
IRP_n69_p6_HC	5	5	5	5	5	5	30	9	6	4	4	4	3	30
IRP_n75_p6_HC	5	5	5	5	5	5	30	11	5	5	4	4	3	32
								10	5	5	5	3	3	31
IRP_n100_p6_HC	6	6	5	5	6	6	34	9	6	5	5	6	3	34
								11	5	5	5	5	3	34

Both tables show that the maximum number of vehicles is used mainly at the initial time periods of the planning horizon. This can be explained (a) by the fact that the inventory level of each customer is equal to zero at the beginning of the planning horizon and (b) by the usage of the backward delivery exchange mutation operator. Actually, the operator satisfies constraint (3), thus avoiding any stock-out. However, it is interesting to observe that with higher inventory holding costs (Table 13), the optimal solution visits customers more frequently. Furthermore, if low-inventory holding costs (Table 14) are applied to the customers, a decrease in the number of times a customer is visited during the planning horizon can be observed since most of the delivery quantities are scheduled at the initial time periods. On the other hand, in the context of the p-VRP, the number of vehicles is nearly the same, as a specific VRP problem should be solved on a daily basis.

Table 16 *Number of Vehicles used During the Planning Horizon (Second Class of Instances)*

Instance	p-VRP							IRP						
	P1	P2	P3	P4	P5	P6	No. of Routes	P1	P2	P3	P4	P5	P6	No. of Routes
IRP_n15_p6_LC	2	2	2	2	2	2	12	4	2	2	1	0	0	9
IRP_n20_p6_LC	2	2	2	2	2	2	12	7	2	1	0	1	0	11
IRP_n22_p6_LC	2	2	2	2	2	2	12	3	2	2	2	1	1	11
IRP_n39_p6_LC	3	3	3	3	3	3	18	8	3	2	1	1	0	15
IRP_n44_p6_LC	3	3	3	3	3	3	18	14	1	1	1	1	0	18
IRP_n50_p6_LC	3	3	3	3	3	3	18	17	1	0	0	0	0	18

To visually verify the above conclusions, the following figures illustrate the solutions of (a) the IRP_n50_p6 (no inventory holding costs), (b) the IRP_n50_p6_HC (high inventory holding cost) and (c) the IRP_n50_p6_LC (low inventory holding cost) benchmark instances. As regards the first solution, 3 routes are scheduled for each day of the planning horizon since the inventory allocation problem is ignored. Concerning the other two solutions, it can be observed that the evolutionary algorithm changes its behavior based on inventory holding cost information. Specifically, the second solution shows that routes are scheduled only for the time periods 1, 3, 4 and 5 (4 of the 6 days). The third solution, due to the low inventory holding costs, indicates the routes that should be scheduled for the time periods 1 and 2 (2 of the 6 days). Therefore, with higher inventory holding costs, the solution visits customers more frequently. Finally, Figure 6 illustrates a typical graph of the minimum IRP cost in the population as a function of generation number. It presents the convergence of fitness values regarding some instances of Class A and B. For each instance, fluctuations can be observed during convergence. However, the entirety direction of evolution indicates improvement with respect to the minimization of inventory routing problem cost. For the majority of instances, it appears that the best candidate solution will continue to improve for a few hundred more generations.

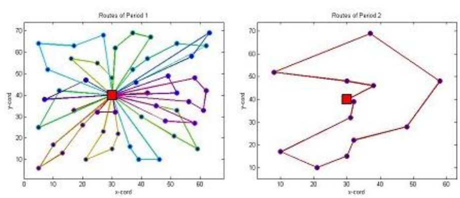


Figure 3 *Solving the IRP with Low Inventory Holding Costs at the Customers (IRP_n50_p6_HC)*

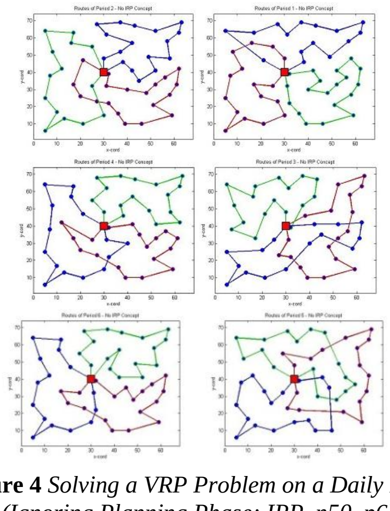


Figure 4 *Solving a VRP Problem on a Daily Basis (Ignoring Planning Phase; IRP_n50_p6)*

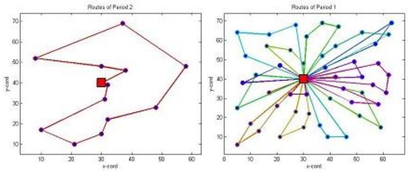


Figure 5 *Solving the IRP with Low Inventory Holding Costs at the Customers (IRP_n50_p6_LC)*

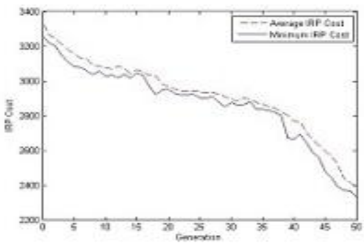
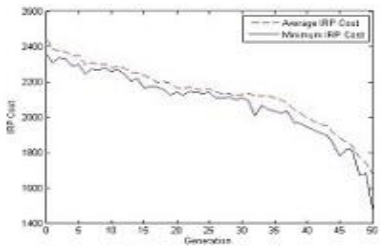
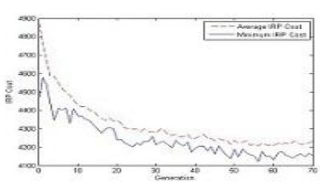
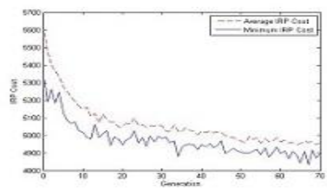
	
IRP_n50_p6_HC(50 Generations)	IRP_n50_p6_LC(50 Generations)
	
IRP_n75_p6_HC(70 Generations)	IRP_n100_p6_HC(70 Generations)

Figure 6 *Convergences of Fitness Values*

6. Discussion and Conclusion

The IRP signifies a challenging stream of research that has drawn the increasing attention of the research community during the last decade. Due to the NP-hard nature of the IRP, this study set out with the aim of assessing the importance of hybrid evolutionary Algorithm in solving the problem. Very little was found in the literature on the combination of population-based search met a heuristics with single-point search met heuristics to handle the IRP. An initial objective of the study was to use a Genetic Algorithm in hybrid synthesis with a Simulated Annealing Algorithm for the solution of the IRP. Particularly, the Genetic Algorithm is related to the planning phase of the hybrid approach to determine the delivery times and quantities, while the Simulated Annealing algorithm is associated with the routing phase to determine the routes of each individual of the population. Both Algorithm are dealt with in an iterative way to define a re-optimization phase. A second objective was to formulate a mathematical problem according to the VMI concept. In this study, stock-outs or lost sales are not allowed, and therefore no shortage costs or costs related to lost sales are included in the objective function. This is a characteristic that differentiates the current study from other works most closely related to this.

Prior studies that have noted the importance of using population-based met a heuristics such as Genetic Algorithm to solve the IRP (e.g., Shukla et al., 2013;

Abdelmaguid and Dessouky, 2006)proposed small IRP instances (5-20 customers) in order to evaluate their Algorithm' performance. The proposed algorithm cannot be compared to other published IRP studies using benchmark instances previously introduced. This is due to the differentiated manner in which the proposed algorithm operates based on the assumptions presented in Section 2 and 3, respectively. Therefore, a third objective of the study was to create new IRP Instances and increase the size of the test problems with reference to the total number of customers. The algorithm has been tested on a newly introduced set of 18 IRP benchmark instances (15-100 customers) by comparing the algorithm's solutions with the solutions obtained by solving a VRP problem for each time period of the planning horizon based on known demand (the planning phase is ignored).

The computational results show that the proposed algorithm is outperformed, simultaneously verifying the benefits obtained by the integration of the inventory and the vehicle routing decisions. These findings further support the idea that in today's economic environment, the ultimate success of a firm depends on its ability to integrate and coordinate supply chain activities (e.g., planning and routing) within the supply chain network. Solutions of the test problems indicate that better solutions are obtained when the planning phase is considered. The ability of each customer to have storage enables a significant decrease in the vehicle routing cost, reducing the total number of routes during the planning horizon and avoiding backorders.

The major contribution of this study is both on the theory of Transportation Science and in the practical implications of vehicle routing and inventory allocation decisions. In terms of the theory, a hybrid evolutionary optimization algorithm for the IRP was proposed. The performance of the algorithm depends on the computational infrastructure (processing power) and the parameters (population size, generations, etc.).Therefore, the algorithm can be further improved by exploring more deeply the parameters of the Genetic Algorithm and the Simulated Annealing Algorithm. In addition, the proposed algorithm can be extended to complicated problems such as the Inventory Routing Problem with Time Windows (IRPTW) and its variations. Both benchmark instances and their solutions are available online. In terms of practice, this study facilitates the development of route planning and inventory routing tools to address real life distribution problems (e.g., distribution of agriculture products).In terms of future research, the goals are (a) to enrich the proposed model taking into account environmental and societal considerations and (b) to enhance the hybrid algorithm so as to be applied in real projects.

7. References

1. Abdelmaguid, T., & Dessouky, M. (2006). A genetic algorithm approach to the integrated inventory-distribution problem. International Journal of Production Research 44(21), 4445-4464.
2. Abdelmaguid, T., Dessouky, M., & Ordóñez, F. (2009). Heuristic approaches for the inventory-routing problem with backlogging. *Computers & Industrial Engineering56* (4), 1519-1534.
3. Adulyasak, Y., Cordeau, J. F., & Jans, R. (2015). The production routing problem: a review of formulations and solution Algorithm. *Computers & Operations Research55*, 141-152.

4. Adulyasak, Y., Cordeau, J.F., & Jans, R. (2014). Formulations and branch-and-cut Algorithm for multivehicle production and inventory routing problems. *INFORMS Journal on Computing 26*(1), 103-120.
5. Aghezzaf, E. H., Raa, B., & Van Landeghem, H. (2006). Modeling inventory routing problems in supply chains of high consumption products. *European Journal of Operational Research 169*(3), 1048-1063.
6. Agra, A., Christiansen, M., & Delgado, A. (2016a). Discrete time and continuous time formulations for a short sea inventory routing problem. *Optimization & Engineering*, doi: 10.1007/s11081-016-9319-0.
7. Agra, A., Christiansen, M., Delgado, A., & Hvattum, L. M. (2015). A maritime inventory routing problem with stochastic sailing and port times. *Computers & Operations Research61*, 18-30.
8. Agra, A., Christiansen, M., Delgado, A., & Simonetti, L. (2014). Hybrid heuristics for a short sea inventory routing problem. *European Journal of Operational Research236* (3), 924-935.
9. Agra, A., Christiansen, M., Ivarsøy, K., Solhaug, I. E., &Tomasgard A. (2016b). Combined ship routing and inventory management in the salmon farming industry. *Annals of Operations Research*, doi: 10.1007/s10479-015-2088-x.
10. Agra, A., Andersson, H., Christiansen, M., & Wolsey, L. (2013). A maritime inventory routing problem: discrete time formulations and valid inequalities. *Networks62* (4), 297-314.
11. Aksen, D., Kaya, O., SibelSalman, F., & Tüncel, O. (2014) An adaptive large neighborhood search algorithm for a selective and periodic inventory routing problem. *European Journal of Operational Research239* (2), 413-426.
12. Andersson, H., Christiansen, M., & Desaulniers, G. (2016). A new decomposition algorithm for a liquefied natural gas inventory routing problem. *International Journal of Production Research 54*(2), 564-578.
13. Andersson, H., Hoff, A., Christiansen, M., Hasle, G., & Løkketangen, A. (2010). Industrial aspects and literature survey: combined inventory management and routing. *Computers & Operations Research37* (9), 1515-1536.
14. Archetti, C., Bertazzi, L., Hertz, A., &Speranza, M. G. (2012). A hybrid heuristic for an inventory routing problem. *INFORMS Journal on Computing24* (1), 101-116.
15. Archetti, C., Boland, N., & Speranza, M. G. (2014). A matheuristic for the multi-vehicle inventory routing problem. Department of Economics and Management, University of Brescia, Italy. Working paper number: WPDEM 2014/3.
16. Archetti, C., Speranza, M. G. (2016). The inventory routing problem: the value of integration. *International Transactions in Operational Research 23*(3), 393-407.
17. Archetti, C., & Speranza, M. G. (2013). A survey on matheuristics for routing problems. Department of Economics and Management, University of Brescia, Italy. Working paper number: WPDEM 2013/11.
18. Arram, A., Ayob, M., & Zakree, M. (2014). Comparative study of meta-heuristic approaches for solving traveling salesman problem. *Asian Journal of Applied Sciences7*, 662-670.

19. Augerat, P., Belenguer, J. M., Benavent, E., Corberán, A., Naddef, D., & Rinaldi, G. (1998). Computational results with a branch-and-cut code for the capacitated vehicle routing problem. Institute for Systems Analysis and Computer Science, Roma, Italy. Research Report number: R.495.
20. Axsäter, S. (2006). *Inventory control*. (2nd ed.). New York: Springer.
21. Aziz, N. A. B., & Moin, N. H. (2007). Genetic algorithm based approach for the multi product multi period inventory routing problem. In *Proceedings of 2007 IEEE international conference on industrial engineering and engineering management*, Singapore (pp. 1619-1623).
22. Ballou, R. H. (1989). Heuristics: rules of thumb for logistics decision making. *Journal of Business Logistics 10* (1), 122-132.
23. Bektas, T. (2006). The multiple traveling salesman problem: an overview of formulations and solution procedures. *Omega34* (3), 209-219.
24. Bell, W. J., Dalberto, L. M., Fisher, M. L., Greenfield, A. J., Jai Kumar, R., Kedia, P., Mack R. G, & Prutzman, P. J. (1983). Improving the distribution of industrial gases with an on-line computerized routing and scheduling optimizer. *Interfaces13* (6), 4-23.
25. Bertazzi, L., Bosco, A., Guerriero, F., & Laganà, D. (2013). A stochastic inventory routing problem with stock-out. *Transportation Research Part C: Emerging Technologies27*, 89-107.
26. Bertazzi, L., Bosco, A., & Laganà, D. (2015). Managing stochastic demand in an inventory routing problem with transportation procurement. *Omega56*, 112-121.
27. Bertazzi, L., Savelsbergh, M., & Speranza, M. G. (2008). Inventory routing. In B. Golden, S. Raghavan, & E. Wasil (Eds.), *the vehicle routing problem: latest advances and new challenges* (pp. 49-72). New York: Springer.
28. Bertazzi, L., & Speranza, M. G. (2013). Inventory routing problems with multiple customers. *EURO Journal on Transportation and Logistics 2*(3), 255-275.
29. Bertazzi, L., & Speranza, M. G. (2012). Inventory routing problems: an introduction. *EURO Journal on Transportation and Logistics1* (4), 307-326.
30. Bräysy, O., & Gendreau, M. (2005a). Vehicle routing problem with time windows, Part I: route construction and local search Algorithm. *Transportation Science39* (1), 104-118.
31. Bräysy, O., & Gendreau, M. (2005b). Vehicle routing problem with time windows, Part II: met a heuristics. *Transportation Science 39*(1), 119-139.
32. Campbell, A. M., & Savelsbergh, M. (2004). A decomposition approach for the inventory-routing problem. *Transportation Science 38*(4), 488-502.
33. Chena, H.K., Hsueh, C.F., & Chang, M.S. (2009). Production scheduling and vehicle routing with time windows for perishable food products. *Computers & Operations Research 36*, 2311-2319.
34. Chen, Z.L. (2004). Integrated production and distribution operations: Taxonomy, models, and review. In D. Simchi-Levi, S.D. Wu, & Z.J. Shen (Eds.), *Handbook of quantitative supply chain analysis: Modeling in the E-business e-ra* (pp. 711-745). New York: Springer.
35. Chen, Z.L., & Vairaktarakis, G. (2005). Integrated scheduling of production and distribution operations. *Management Science 51*(4), 614-628.

36. Cho, W. D., Lee, Y. H., Lee, Y. T., & Gen, M. (2013). An adaptive genetic algorithm for the time dependent inventory routing problem. *Journal of Intelligent Manufacturing25* (5), 1025-1042.
37. Coelho, L. C., Cordeau, J. F., & Laporte, G. (2013). Thirty years of inventory routing. *Transportation Science 48*(1), 1-19.
38. Coelho, L., Cordeau, J. F., & Laporte G. (2012a). The inventory-routing problem with transhipment. *Computers & Operations Research39* (11), 2537-2548.
39. Coelho, L., Cordeau, J. F., & Laporte, G. (2012b). Consistency in multi-vehicle inventory-routing. *Transportation Research Part C: Emerging Technologies24*, 270-287.
40. Coelho, L. C., & Laporte G. (2013). The exact solution of several classes of inventory-routing problems. *Computers & Operations Research 40*(2), 558-565.
41. Cordeau, J. F., Gendreau, M., Laporte, G., Potvin, J. Y., & Semet, F. (2002). A guide to vehicle routing heuristics. *The Journal of the Operational Research Society 53*(5), 512-522.
42. Croom, S., Romano, P., & Giannakis, M. (2000). Supply chain management: an analytical framework for critical literature review. *European Journal of Purchasing & Supply Management 6*(1), 67-83.
43. Dantzig, G. B., & Ramser, J. H. (1959). The truck dispatching problem. *Management Science 6*(1), 80-91.
44. Diabat, A., Abdallah, T., & Le, T. (2016). A hybrid tabu search based heuristic for the periodic distribution inventory problem with perishable goods. *Annals of Operations Research242* (2); 373-398.
45. Díaz-Parra, O., Ruiz-Vanoye, J., Loranca, B. B., Fuentes-Penna, A., & Barrera-Cámara, R. (2014). A survey of transportation problems. *Journal of Applied Mathematics*, http://dx.doi.org/10.1155/2014/848129.
46. Eksioglu, B., Vural, A. V., & Reisman, A. (2009). The vehicle routing problem: a taxonomic review. *Computers & Industrial Engineering 57*(4), 1472-1483.
47. El-Sherbeny, N. (2010). Vehicle routing with time windows: an overview of exact, heuristic and met a heuristic methods. *Journal of King Saud University 22*(3), 123-131.
48. Federgruen, A., Prastacos, G., & Zipkin, P. H. (1986). An allocation and distribution model for perishable products. *Operations Research 34*(1), 75-82.
49. Federgruen, A., & Zipkin, P. H. (1984). A combined vehicle-routing and inventory allocation problem. *Operations Research 32*(5), 1019-1037.
50. Flood, M. (1956). The traveling-salesman problem. *Operations Research 4*(1), 61-75.
51. Gaur, V., & Fisher, M. L. (2004). A periodic inventory routing problem at a supermarket chain. *Operations Research 52*(6), 813-822.
52. Ghiami, Y., Van Woensel, T., Christiansen, M., & Laporte G. (2015). A combined liquefied natural gas routing and deteriorating inventory management problem. In F. Corman, S. Voβ, & R. Negenborn (Eds.), *Computational Logistics* (pp. 91-104). Switzerland: Springer.
53. Goel, V., Furman, K., Song, J. H., & El-Bakry, A. (2012). Large neighborhood search for LNG inventory routing. *Journal of Heuristics18* (6), 821-848.

54. Goel, V., Slusky, M., Van Hoeve, W. J., Furman, K. C., & Shao, Y. (2015). Constraint programming for LNG ship scheduling and inventory management. *European Journal of Operational Research 241*(3), 662-673.
55. Goetschalckx, M. (2011). *Supply chain engineering*. New York: Springer.
56. Griffis, S. E., Bell, J. E., & Closs, D. J. (2012). Met a heuristics in logistics and supply chain management. *Journal of Business Logistics 33*(2), 90-106.
57. Guemri, O., Bektar, A., Beldjilali, B., &Trentesaux, D. (2016). GRASP-based heuristic algorithm for the multi-vehicle inventory routing problem. *4OR14* (4); 377-404.
58. Guerrero, W. J., Prodhon, C., Velasco, N., & Amaya, C. A. (2013). Hybrid heuristic for the inventory location-routing problem with deterministic demand. *International Journal of Production Economics 146*(1), 359-370.
59. Hemmati, A., Hvattum, L. M., Christiansen, M., & Laporte, G. (2016). An iterative two-phase hybrid mat heuristic for a multi-product short sea inventory-routing problem. *European Journal of Operational Research252* (3), 775-788.
60. Hemmati, A., Stålhane, M., Hvattum, L. M., & Andersson, H. (2015). An effective heuristic for solving a combined cargo and inventory routing problem in tramp shipping. *Computers & Operations Research 64*, 274-282.
61. Hewitt, M., Nemhauser, G., Savelsbergh, M., & Song, J. H. (2013). A branch-and-price guided search approach to maritime inventory routing. *Computers & Operations Research 40*(5), 1410-1419.
62. Holland, J. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press.
63. Huang, S. H., & Lin, P. C. (2010). A modified ant colony optimization algorithm for multi-item inventory routing problems with demand uncertainty. *Transportation Research Part E: Logistics and Transportation Review 46*(5), 598-611.
64. Jiang, Y., & Grossmann, I. (2015). Alternative mixed-integer linear programming models of a maritime inventory routing problem. *Computers & Chemical Engineering 77*, 147-161.
65. Juan, A., Grasman, S, Caceres-Cruz J., & Bektaş T. (2014). A simheuristic algorithm for the single-period stochastic inventory-routing problem with stock-outs. *Simulation Modelling Practice and Theory 46*, 40-52.
66. Kirkpatrick, S., Gellat, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science 220*(4598), 671-680.
67. Laporte, G. (2010). A concise guide to the traveling salesman problem. *The Journal of the Operational Research Society61* (1), 35-40.
68. Laporte, G. (2009). Fifty years of vehicle routing. *Transportation Science 43*(4), 408-416.
69. Lappas, P., Kritikos, M., & Ioannou, G. (2015a). A genetic algorithm for the inventory routing problem with time windows. *27th European conference on operational research*, Glasgow, United Kingdom.
70. Lappas, P., Kritikos, M., & Ioannou, G. (2015b). Classic met a heuristics and evolutionary optimization Algorithm for routing problems: a computational study. *4th International Eurasian Conference on Mathematical Sciences and Applications*, Athens, Greece.

71. Lappas, P., Kritikos, M., & Ioannou, G. (2015c). Met a heuristic algorithm for routing problems: from the travelling salesman problem to the inventory routing problem. *MASSEE International Congress on Mathematics*, Athens, Greece.
72. Le, T., Diabat, A., Richard, J. P., & Yih Y. (2013). A column generation-based heuristic algorithm for an inventory routing problem with perishable goods. *Optimization Letters 7*(7), 1481-1502.
73. Li, K., Chen, B., Lyer Siva kumar, A., & Wu, Y. (2014). An inventory-routing problem with the objective of travel time minimization. *European Journal of Operational Research 236*(3), 936-945.
74. Li, Z., Jiang, C., & Jiang, L. (2015). An inventory routing problem with soft time windows. *12th International Symposium on Operations Research and its Application in Engineering, Technology and Management*, Luoyang, China.
75. Liao, L., Li, J., & Wu, Y. (2013). Modeling and optimization of inventory-distribution routing problem for agriculture products supply chain. *Discrete Dynamics in Nature and Society*, http://dx.doi.org/10.1155/2013/409869.
76. Liu, S. C., & Lee, W. T. (2011). A Heuristic method for the inventory routing problem with time windows. *Expert Systems with Applications 38*(10), 13223-13231.
77. Liu, S., Lei, L., & Ruszczynski, A. (2007). Production and distribution planning with many practical constraints. *International Journal of Operations and Quantitative Management 13*(2), 129-143.
78. Liu, S. C., Lu, M. C., & Chung, C. H. (2015). A hybrid heuristic method for the periodic inventory routing problem. *The International Journal of Advanced Manufacturing Technology85* (9), 2345-2352.
79. Mabrouk, N., Eddaly, J., & Chabchoub, E. (2010). A heuristic approach for the integrated production and distribution problem. *8th International Conference on Modeling and Simulation*, Hammamet, Tunisia.
80. Maniezzo, V., Stützle, T., & Voβ, S. (2009). *Matheuristics: hybridizing metaheuristics and mathematical programming*. New York: Springer.
81. Melachrinoudis, E., & Min, H. (2009). Design of a two-level distribution network using the Voronoi diagram. *International Journal of Operations and Quantitative Management 15*(1), 29-44.
82. Mercer, A., & Tao, X. (1996). Alternative inventory and distribution policies of a food manufacturer. *The Journal of the Operational Research Society 47*(6), 755-765.
83. Min, H., & Zhou, G. (2002). Supply chain modeling: past, present and future. *Computers & Industrial Engineering 43*(1-2), 231-249.
84. Mirzaei, S., & Seifi, A. (2015). Considering lost sale in inventory routing problems for perishable goods. *Computers & Industrial Engineering 87*, 213-227.
85. Mjirda, A., Jarboui, B., Macedo, R., & Hanafi, S. (2012). A variable neighborhood search for the multi-product inventory routing problem. *Electronic Notes in Discrete Mathematics 39*, 91-98.
86. Mjirda, A., Jarboui, B., Macedo, R., Hanafi, S., & Mladenović, N. (2014). A two phase variable neighborhood search for the multi-product inventory routing problem. *Computers & Operations Research52*, 291-299.

87. Mjirda, A., Jarboui, B., Mladenović, J., Wilbaut, C., & Hanafi, S. (2016). A general variable neighborhood search for the multi-product inventory routing problem. *IMA Journal of Management Mathematics 27*, 39-54.
88. Moin, N. H., & Salhi, S. (2007). Inventory routing problems: a logistical overview. *The Journal of the Operational Research Society 58*(9), 1185-1194.
89. Moin, N. H., Salhi, S., & Aziz, N. A. B. (2011). An efficient hybrid genetic algorithm for the multi-product multi-period inventory routing problem. *International Journal of Production Economics 133*(1), 334-343.
90. Nambirajan, R., Mendoza, A., Pazhani, S., Narendran, T. T., & Ganesh, K. (2016). CARE: heuristics for two-stage multi-product inventory routing problems with replenishments. *Computers & Industrial Engineering 97*, 41-57.
91. Neubert, G., Savino, M., & Pedicini, C. (2010). Simulation approach to optimize production costs through value stream mapping. *International Journal of Operations and Quantitative Management 16*(1), 1-21.
92. Niakan, F., & Rahimi, M. (2015). A multi-objective healthcare inventory routing problem; a fuzzy possibilistic approach. *Transportation Research Part E: Logistics and Transportation Review80*, 74-94.
93. Nikzad, P., Min, H., & Bozorgi-Amiri, A. (2017). Risk pooling strategy for redesigning multi-echelon distribution networks. *International Journal of Operations and Quantitative Management 23*(3), 187-210.
94. Nolz, P., Absi, N., & Feillet, D. (2014a). A stochastic inventory routing problem for infectious medical waste collection. *Networks 63*(1), 82-95.
95. Nolz, P., Absi, N., & Feillet, D. (2014b). A bi-objective inventory routing problem for sustainable waste management under uncertainty. *Journal of Multi-criteria Decision Analysis21* (5-6), 299-314.
96. Papageorgiou, D., Nemhauser, G., Sokol, J., Cheon, M. S., & Keha, A. (2014). MIRPLib – a library of maritime inventory routing problem instances: survey, core model, and benchmark results. *European Journal of Operational Research 235*(2), 350-366.
97. Park, Y. B., Yoo J. S., & Park, H. S. (2016). A genetic algorithm for the vendor-managed inventory routing problem with lost sales. *Expert Systems with Applications 53*, 149-159.
98. Popović, D., Vidović, M., & Radivojević, G. (2012). Variable neighborhood search heuristic for the inventory routing problem in fuel delivery. *Expert Systems with Applications 39*(18), 13390-13398.
99. Potvin, J. Y. (2009). Evolutionary Algorithm for vehicle routing. *INFORMS Journal on Computing21* (4), 518-548.
100. Qin, L., Miao, L., Ruan, Q., & Zhang, Y. (2014). A local search method for periodic inventory routing problem. *Expert Systems with Applications41* (2), 765-778.
101. Raa B. (2015). Fleet optimization for cyclic inventory routing problems. *International Journal of Production Economics160*, 172-181.
102. Rego, C., Gamboa, D., Glover, F., & Osterman, C. (2011). Traveling salesman problem heuristics: leading methods, implementations and latest advances. *European Journal of Operational Research 211*(3), 427-441.
103. Ronen, D. (1993). Ship scheduling: the last decade. *European Journal of Operational Research 71*(3), 325-333.

104. Santos, E., Satoru Ochi, L., Simonetti, L., & Henrique Gonsález, P. (2016). A hybrid heuristic based on iterated local search for multivehicle inventory routing problem. *Electronic Notes in Discrete Mathematics 52*, 197-204.
105. Savino, M., Brun, A., & Mazza, A. (2014). Dynamic workforce allocation in a constrained flow shop with multi-agent system. *Computers in Industry 65*, 967-975.
106. Savino M, & Mazza, A. (2015). Kanban-driven parts feeding within a semi automated O-shaped assembly line: A case study in the automotive industry. *Assembly Automation 35*(1, 2), 3-15.
107. Savino, M., Mazza, A., & Neubert, G. (2014). Agent-based flow-shop modeling in dynamic environment. *Production Planning & Control 25*(2), 110-122.
108. Schmid, V., Doerner, K., & Laporte, G. (2013). Rich routing problems arising in supply chain management. *European Journal of Operational Research224* (3), 435-448.
109. Shao, Y., Furman, K., Goel, V., & Hoda, S. (2015). A hybrid heuristic strategy for liquefied natural gas inventory routing. *Transportation Research Part C: Emerging Technologies53*, 151-171.
110. Shirokikh, V., & Zakharov, V. (2015). Dynamic adaptive large neighborhood search for inventory routing problem. In H. A. L. Thi, T. P. Dinh, & N. T. Nguyen (Eds.), *Modeling, computation and optimization in information systems and management sciences* (pp. 231-241). New York: Springer.
111. Shukla, N., Tiwari, M. K., & Ceglarek, D. (2013). Genetic-Algorithm-based algorithm portfolio for inventory routing problem with stochastic demand. *International Journal of Production Research51* (1), 118-137.
112. Simić, D., &Simić, S. (2013). Evolutionary approach in inventory routing problem. In I. Rojas, G. Joya, & J. Cabestany (Eds.), *Advances in computational intelligence* (pp. 395-403). New York: Springer.
113. Singh, T., Arbogast, J., & Neagu, N. (2015). An incremental approach using local-search heuristic for inventory routing problem in industrial gases. *Computers & Chemical Engineering 80*, 199-210.
114. Song, J. H., & Furman, K. (2013). A maritime inventory routing problem: practical approach. *Computers & Operations Research 40*(3), 657-665.
115. Soysal, M., Bloemhof-Ruwaard, J., Haijema, R., & Van der Vorst, J. (2016). Modeling a green inventory routing problem for perishable products with horizontal collaboration. *Computers & Operations Research*, http://dx.doi.org/10.1016/j.cor.2016.02.003.
116. Soysal, M., Bloemhof-Ruwaard, J., Haijema, R., Van der Vorst, J. (2015). Modeling an inventory routing problem for perishable products with environmental considerations and demand uncertainty. *International Journal of Production Economics164*, 118-133.
117. Staggemeier, A., Clark, A., Aickelin, U., & Smith, J. (2002). A hybrid genetic algorithm to solve a lot-sizing and scheduling problem. *16th Triannual Conference of the International Federation of Operational Research Societies*, Edinburgh, Scotland.
118. Sujan, P., Katsuhiko, T., & Katsumi, M. (2015). Simultaneous quotations with capacity planning under contingent orders. *International Journal of Operations and Quantitative Management 21*(2), 99-125.

119. Talbi, E. G. (2009). *Metaheuristics: from design to implementation*. New Jersey: Willey.
120. Tan, K. C. (2001). A framework of supply chain management literature. *European Journal of Purchasing and Supply Management 7*(1), 39-48.
121. Tatsis, V., Parsopoulos, K., Skouri, K., & Konstantaras, I. (2013). An ant-based optimization approach for inventory routing. In M. Emmerich, A. Deutz, O. Schütze, T. Bäck, E. Tantar, A. Tantar, P. Del Moral, P. Legrand, P. Bouvry, & C. Coello (Eds.), *EVOLVE – A bridge between probability, set oriented numerics, and evolutionary computation IV* (pp. 107-121). New York: Springer.
122. Toth, P., & Vigo, D. (2002). *The vehicle routing problem*. Philadephia: SIAM.
123. Vansteenwegen, P., & Mateo, M. (2014). An iterated local search algorithm for the single-vehicle cyclic inventory routing problem. *European Journal of Operational Research 237*(3), 802-813.
124. Yang, Z., Emmerich, M., Bäck, T., & Kok, J. (2015). Multicriteria inventory routing by cooperative swarms and evolutionary Algorithm. In J. M. Ferrández Vicente, J. R. Álvarez-Sánchez, F. de la Paz López, J. Toledo-Moreo, & H. Adeli (Eds.), *Bioinspired computation in artificial systems* (pp. 127-137). New York: Springer.
125. Zeng, W., & Zhao, Q. (2010). Study of stochastic demand inventory routing problem with soft time windows based on MDP. In Z. Zeng, & J. Wang (Eds.), *Advances in neural network research and applications* (pp. 193-200). Shanghai: Springer.
126. Zhang, C., Nemhauser, G., Sokol, J., Cheon, M. S., & Keha, A. (2013). Flexible solutions to maritime inventory routing problems with delivery time windows. Georgia Institute of Technology, Technical Report.
127. Zhang, C., Nemhauser, G., Sokol, J., Cheon, M. S., & Papageorgiou, D. (2015). Robust inventory routing with flexible time window allocation. Georgia Institute of Technology, Technical Report.

About Our Authors

Pantelis Z. Lappas has completed his doctoral studies in Operations Research at Athens University of Economics and Business (2017). He holds a MSc degree in Information Systems with specialization in Business Information Systems (2009) and a BSc degree in Informatics Engineering (2007). He has also completed a MSc degree in Statistics with specialization in Quantitative Statistical Methods applied in Decision Making (2013). In addition, he attended 1-Year Pedagogical Training Program acquiring pedagogical and didactic competence (2011). His research interests are rooted in Operations Research. He is primarily interested in Stochastic Inventory Routing Problems, Queuing Networks and Evolutionary Algorithm.

Manolis N. Kritikos is Assistant Professor of Operations Research and Information Systems at the Department of Management Science and Technology of Athens University of Economics and Business (AUEB). He received a BSc in Mathematics and a MSc in Computer Science and Operations Research, both from the University of Athens. He holds a PhD in Computer Science and Management from AUEB. His research interests are in the area of Combinatorial Optimization, Logistics, Design and Analysis of Algorithm, Heuristics, Meta-heuristics, Geographical Information

Systems, Data Envelopment Analysis, Mathematics and novel solution methods for the various versions of the Vehicle Routing Problem.

George D. Ioannou is Professor of Production and Operations Management at Athens University of Economics and Business. He received his Diploma in Mechanical Engineering from the National Technical University of Athens, and his MSc/DIC in Industrial Robotics and Manufacturing Automation from Imperial College, London, UK. He received his PhD in Mechanical Engineering from the University of Maryland at College Park, USA. His research concentrates on the quantitative and analytical study of business systems, and merges operations research tools with modern information technology to address open problems faced by today's complex enterprises and supply chain networks.

Optimal Combination of Renewable Energies for an Enterprise–A Case of Taiwan

Hsiao-Fan Wang
Pei-Wen Chou
National Tsing Hua University
(hsiaofanwang@gmail.com)
(suenodequimera@gmail.com)

Volume 24, *Number* 2
June 2018, *pp.* 117-138

To achieve the goals of international carbon reduction commitment without restricting electricity usage and with a valid electricity price, an energy policy was implemented by the government of Taiwan in November 2011. This study proposes a two-stage analytical model to support private enterprises in evaluating the possible generation combination of renewable energy while considering the intermittence of renewable energy, environmental protection and generation cost, and given the monthly electricity demand. The model involves a parameter evaluation model developed to determine the capacity factor of renewable energies. A multi-objective model is also proposed to optimize the mixed generation of renewable energies. A case study of optimal mixed-generation of renewable energies was carried out at the National Tsing Hua University to illustrate and verify the proposed system.

Keywords: Renewable Energy, Generation Combination, Multi-Objective Model, Intermittence, Parameter Evaluation, NTHU

1. Introduction

Growing concern for climate change and nuclear safety has shifted attention on renewable energy as an alternative for electricity generation systems. With the depletion and price volatility of fossil fuels, Taiwan, which is highly dependent on imported fossil fuels, has yet to develop its own renewable power plant. On July 27, 2015, the Taiwanese government set a goal of 30.70% power system device capacity from renewable energy by 2030. Hence, to satisfy this goal, the government has developed two of the largest installation schemes for onshore wind and solar power, with targets set to reach 12.00×10^2 MW and 87.00×10^2 MW by 2030, respectively (Bureau of Energy, Ministry of Economic Affairs, 2015).

Social responsibility and environmental protection has caused many enterprises to exert efforts in reducing the carbon emission intensity of their products, enhance the competitiveness of their export products, and realize their social responsibility by subscribing to green electricity. In addition, if enterprises can develop their own green power generation, they can promote their corporate image further and enhance their intangible assets. However, several limitations due to economic, weather, and environment issues need to be addressed before we can proceed with renewable energy penetration. These limitations will be discussed in the next sections.

Rom the economic aspect, two factors raised the cost of renewable energy generation. One factor is the cost from the factory setting and operation maintenance, and the other is the cost of carbon emission. Taiwan has formulated government policies that provide subsidies for those who want to develop renewable energy systems, mainly in the form of Feed-in-Tariffs. The government has also issued a policy committing to purchase all electricity from private generation with prices higher than the market. With the improvement of generation performance and low-carbon technologies, the power system will secure future energy supplies and mitigate its effects on the environment. Moreover, investment cost will be decreased due to advances in technology, which is a main force in the future development of renewable energy generation systems (International Energy Agency, 2015).

Taiwan has been actively participating in international affairs, and has in fact taken the initiative to announce the Intended Nationally Determined Contribution (INDC) with the hopes of leading its people in building a low-carbon environment. With the Greenhouse Gases Reduction and Management Act promulgated in July 2015, Taiwan INDC's target for the 2030 reduction level can be regarded as a goal of 50% reduction below the 2005 level by 2050 (Environment Protection Administration Executive Yuan, Taiwan, 2015).

Wind and solar photovoltaic power are both abundant and relatively available in Taiwan, and hence, most renewable electricity supply will be based on these two resources. However, the restriction of renewable energies is that they are highly dependent on weather conditions and are intermittent due to seasonal variations. Seasonal variation and uncertainty of daily sunlight or wind speed required in the operation of power system taking the capacities of each resource they need to install into consideration.

Therefore, in this study we intend to decide not only on the renewable energy supply mix but also on how much electricity the enterprise should purchase monthly from the major power company.

The Executive Yuan encourages government agencies and schools to reduce carbon emission. On November 17, 2014, it spearheaded a program that encouraged saving power, fuel, water, and paper to increase energy conservation. It is anticipated that government agencies and schools can provide guidance to civil society to adopt energy saving schemes (Bureau of Energy, Ministry of Economic Affairs, 2014). Therefore, the National Tsing Hua University (NTHU) of Taiwan will be taken as our case study. The NTHU has an available roof area of 79.61×10^3 square meters, which will serve as our land resource for the renewable power plant to implement comprehensive carbon reduction.

2. Literature Review

This chapter reviews the related literature on energy management. The methods used for evaluation and to implement the production plan for renewable energy are presented in section 2.2. The related models of energy supply planning are discussed in section 2.3. Finally, the summary and conclusion are presented in section 2.4.

2.1 Energy Management

Conventional power generation system focuses on the mix of non-renewable and renewable energies; however, the price of raw materials for non-renewable energies is unstable. For instance, most fossil fuel resources are often located in politically sensitive areas of world. Hence, when decision makers plan a power generation system, price volatility is the factor that can be difficult to control (Ryu et al., 2014).

With the shortcomings of non-renewable energies, the Taiwanese government seeks to increase renewable energy penetration in Taiwan. However, because the investment cost of renewable energies system is too enormous for an enterprise, the government provides subsidies under various situations to promote the penetration of renewable energy. Currently, major energy producers and consumers, such as the United States, Germany, China, Japan and Brazil have also formulated policies to support renewable energy production with subsidies (Zhang et al., 2014).

In recent years, China has rapidly developed wind power, solar photovoltaic, and other renewable energies; however, despite the development of the renewable energy industry, generating electricity from merged grids can be difficult. Hence, to address the above problem and to promote the development and utilization of renewable energy, a development fund for renewable energy was established in 2012. The subsidized projects include wind power, biomass power, solar power, geothermal power, and ocean energy power generation. Moreover, the power plant can include investment, operation, and maintenance costs, which are connected to the grid project, to the subsidy (State Grid Energy Research Institute, 2012).

Regardless of the subsidy programs established in different countries, all of these countries are moving toward the same goal, that of renewable energy penetration. The rise of environmental issues and volatility of fuel prices will eventually push countries around the world to focus more attention on sustainable development.

2.1.1 Renewable Energies Development of Enterprise

Electricity consumption by industries is always larger than that by households. Therefore, industries are very sensitive to fluctuations in electricity prices. The Taiwanese government has set a reduction target of carbon emissions through alternative generation plans, which has raised concerns for enterprises in terms of higher electricity prices in the future. The total electricity costs can only be controlled by demand-side management for an enterprise because only one Power Company is non-private in Taiwan. Therefore, enterprises need to develop their own power plants to reduce uncertainty caused by energy depletion and price volatility.

Park and Kwon (2016) use simulation to explore renewable power generation systems for Kyung-Hee University's Global Campus in South Korea. They use a generating system with different kinds of components, such as PV panels, diesel generators, wind turbines, batteries, and grid to conduct a scenario analysis. They determined the optimal energy mix that could meet electricity demands at a minimum cost. However, their study has some limitations. First, the study did not discuss renewable portfolio standard policies and environment protection operated by the South Korean government. Second, they did not discuss the economic scales and learning effects. Finally, because their study is based on a university in South Korea, their conclusions cannot be adopted to other regions because of the different climate effects in different regions.

Pineda and Bock (2016) design a game for two players, one of which is a renewable energy generating company that will decide on installed capacity and capacity factor, while the other is a non-renewable energy generating company with known fuel cost and fixed capacity. The renewable energy generating company can sell electricity and green certificates that can be used to comply with the renewable quota. To avoid the non-compliance penalty, the non-renewable energy generating company has to buy green electricity from the renewable energy generating company.

The result of this study shows that with the exception of the quota obligation, the non-compliance penalty has a significant effect on investments in renewable energy-based generation. They also determine the capacity of the renewable energy generating company by maximizing social welfare, and the effect of the level of competition on the decisions of the renewable energy generating company. Finally, they point out that the Feed-in-Tariff subsidy or sensitivity analysis tool can be considered in future studies.

In developing a renewable energy generation system, we need to consider many external and internal factors for an enterprise. In terms of current trends, many companies are beginning to focus on social responsibility. Social responsibility is a company's intangible asset that can bring potential benefits. This study will focus on the situation in Taiwan by providing a model for an enterprise to plan a renewable energy generation system that considers the costs, emissions, and government policies.

2.1.2 Environmental Protection

With the increase in greenhouse gases, CO_2 emission has become the real cause of climate change. Suspended particles of CO_2 and SO_2 cannot be eliminated by current cleaning equipment. Therefore, a carbon tax must be promulgated to reduce greenhouse gas emissions and enable people to conserve energy. Carbon emissions from fossil fuels are generated from mining to the power generation process and its final disposal. Carbon emission can be managed and controlled efficiently if we can examine the carbon content at each stage, from mining to disposal. Examining the carbon content will also serve as a reference for moving ahead in carbon capture and storage technology at each stage. Most fossil fuels are imported from other countries, and hence, carbon emission from mining is not considered in Taiwan. However, the price of fossil fuels may be affected by the carbon tax in the country from which it is imported. Chen et al. (2015) investigate the effects of carbon tax on imported forest products and illustrate that carbon tax causes the prices of imported harvested wood products to rise and the quantity of imported harvested wood products to decline. Consequently, levying a carbon tax should not only include direct carbon emissions, but also the price of imported products. As stated earlier, moving forward in our environmental protection goals requires that we consider the carbon emissions of a power generation system from start to finish.

The Cost Assessment for Sustainable Energy Systems (CASES) is an EU research project that studies the effects of various power generation technologies on the environment and quantifies the cost (Nuclear Energy Agency, 2012). Stringent estimates are necessary to provide quantitative analysis for decision-makers. The CASES is an assessment of the combination of the effects of climate change and human health risks. The study shows that the burning of fossil fuels must withstand high external costs. Air pollution caused by burning biomass is often harmful to people. In contrast, pollution from solar power generation is rare, but the damage to human health is much higher than hydro, wind, and nuclear power as shown in Table 1.

Several countries which have implemented the carbon tax policy include Norway, Sweden, and the United Kingdom. Although carbon tax has not been implemented in Taiwan, we could refer to the experiences of other countries to evaluate environment costs. The greenhouse gas emissions of different countries are shown in Table 2.

Table 1 *The Hazard Costs from 2005 to 2010 in Europe*
(Source: Nuclear Energy Agency, 2012)

External Cost ($NT/KWh)	Nuclear Energy	Coal Fired IGCC	Lignite IGCC	Gas CCGT	Hydro	onshore	offshore	solar	Biomass (Herbal)	Biomass (Woody)
Damage to human health	0.062	0.334	0.154	0.170	0.023	0.030	0.029	0.263	0.622	0.186
Biodiversity loss	0.004	0.032	0.013	0.021	0.001	0.002	0.001	0.014	0.118	0.020
Agricultural damage (N,O_3,SO_2)	0.004	0.032	0.013	0.021	0.001	0.002	0.001	0.014	0.118	0.020
Air pollution (SO_2,NO_2)	0.001	0.004	0.001	0.003	0.000	0.000	0.000	0.004	0.005	0.003
Radioactive nuclides	0.001	0	0	0	0	0	0	0	0	0
Climate Change	0.017	0.702	0.783	0.359	0.006	0.008	0.007	0.072	0.058	0.048
Total cost	0.086	1.078	0.952	0.557	0.030	0.041	0.038	0.355	0.807	0.262

Table 2 *The Main Emission Index of Different Countries*
(Source: International Energy Agency, 2013, Key World Energy Statistics 2013)

	CO_2 **(Million Tons)**	**Population (Million)**	CO_2 **/Population (Tons/Capita)**	CO_2 **/GDP(PPP) (kg/USD)**
Taiwan	26.47×10^{1}	23.39	11.31	34.00×10^{-2}
Japan	11.86×10^{2}	12.78×10^{1}	92.80×10^{-1}	30.00×10^{-2}
Korea	58.77×10^{1}	49.78	11.81	43.00×10^{-2}
United States	52.87×10^{2}	31.20×10^{1}	16.94	40.00×10^{-2}
China	79.55×10^{2}	13.44×10^{2}	59.20×10^{-1}	34.00×10^{-2}

Table 2 shows that regardless of which country has implemented a carbon trading policy whether based on the population or the average carbon emission per person, the South Korean experience is the closest to Taiwan. The industrial structure of South Korea is also similar to Taiwan. Therefore, we choose South Korea as a reference for carbon tax and carbon trading policy in the measurement of environment cost.

2.2 Evaluation Method of Renewable Energy

Wind and solar energies are both abundant and relatively available in Taiwan, making it easier for enterprises to develop electricity plants using these types of energies. Therefore, the most possible renewable electricity supply will be based on these two resources. However, a restriction of renewable energies is that their supply is intermittent. Seasonal variations and uncertainty of daily sunlight or wind speed require that we take the capacity factor of renewable energy into consideration.

Historical data on the capacity factor differ in different areas. For wind power in particular, data differ because most large-scale wind turbines are located along the seashore. However, based on the technology development, small-scale wind turbines have been commonly established in urban area. In this study, we hope to come up with a capacity factor for enterprises in general regardless of location. Therefore, in this study, we will develop an evaluation model on wind speed and the amount of sunlight to enhance the use of renewable energy.

1. Capacity Factors of Solar Energy: The amount of sunlight is affected by the sun, weather, season, and temperature of a region. Usually, the daily global irradiance on the horizontal plane is used to represent the amount of sunlight. The most commonly used probability distribution function of sunlight includes Log-normal distribution, Weibull distribution, and Beta distribution (Shen, 2000).

In addition to the frequency distribution of sunlight, the performance ratio, which is decided by its installation and the quality of its components, are marked in current versions of solar panels. With the advancement of technology, using the performance ratio to estimate the capacity factor of solar power can be more realistic and more applicable for the study (Filippo et al., 2013).

2. Capacity Factors of Wind Energy: A number of density functions have been used to describe the wind speed frequency curve, the most famous among which are the Weibull and Rayleigh functions. The Weber function is a general gamma function that contains two parameters, whereas Rayleigh is a subset of Weber distribution and contains one parameter. As a result, the Weibull distribution function can be used more widely than the Rayleigh distribution function (Ragab et al., 1993).

A number of probability distribution of wind speed has been proposed. In particular, the Weibull probability distribution, which is a well-known and widely used in various wind energy related investigations, has been found to provide good performance within a short period. The accuracy of the Weibull distribution has also been confirmed by Arslan et al. (2014), Islam et al. (2011), and Kwon (2010). We will choose the two-parameter Weibull distribution function as our assessment method to meet the monthly time intervals.

2.3 Multi-objective Method

The difficulty of renewable energy penetration is its intermittence and instability. Moura and Almeida (2010) developed a multi-objective optimization for power generation system to achieve the minimal renewable energies intermittence, production cost, and peak load. They use the method of shortening the gap between the capacity factor and the past monthly average capacity factor to reduce variations in generation, which allowed them to narrow the intermittence of renewable energy. Sharafi and ELMekkawy (2014) propose a novel approach for hybrid renewable energy system, including various generators and storage devices while minimizing the total cost of the system, unmet load, and fuel emission. Bilil et al. (2014) present a multi-objective formulation to optimize both the annualized renewable energy cost and the system reliability simultaneously.

Ho et al. (2014) propose a multi-objective linear programming and fuzzy two-stage algorithm to reduce carbon emissions. Therefore, in the case of multiple targets, we can provide different weights to select an appropriate multi-objective programming according to different objects (Cheng et al., 2013). Numerous studies have used multi-objective programming to investigate the development of power systems, and most studies consider generation cost, carbon emissions, and intermittent issues.

2.4 Discussion and Conclusion

A variety of renewable power supply system can reduce the risk of price changes in fossil fuel and carbon emissions. A number of studies have proposed multi-objective programming to study the energy supply mix. However, the difference in this study is that we intend to decide on the combination of renewable energies and achieve a monthly balance between supply and demand for an enterprise. Hence, the results of this study will include not only the renewable energy supply mix but also how much electricity the enterprise should purchase from the power company for each month. Consequently, we will develop a multi-objective model to achieve the goals economic, environmental, and sustainable optimization for enterprises.

3. Model Formulation

3.1 Problem Statement

We divide the mathematical model into two levels to achieve the goals of cost minimization, carbon reduction, and sustainable power supply. First, because the weather changes over time, a probability distribution function is used to describe the characteristics of the weather such that the capacity factor, which is affected by the weather, can be evaluated monthly using the proposed parameter evaluation model. Consequently, the capacity factors of renewable energies that are most likely to generate power generations can be determined.

Second, by working on the three goals, a multi-objective analytical model is developed to decide on how much of the demand should be provided by the renewable energy enterprise itself and how much should be purchased from the electricity company. Therefore, under this multi-objective analytical model, while the total energy generation of renewable energies is maximized, carbon emission from manufacture will be minimized at minimum cost. The monthly time interval will be adopted to indicate seasonal variations and intermittence of renewable energy.

3.2 Framework of Multi-objective Programming with Parameter Evaluation Model

The framework of the proposed multi-objective programming with parameter evaluation model is illustrated in Figure 1. Input data include monthly average wind speed, monthly amount of sunlight, generation cost per unit of electricity, purchase price per unit of electricity, environmental cost per unit of electricity, and monthly power demand. We propose a parameter evaluation model to determine the capacity factors of renewable energies. Power generation of renewable energies will be maximized under minimal environment and generation costs.

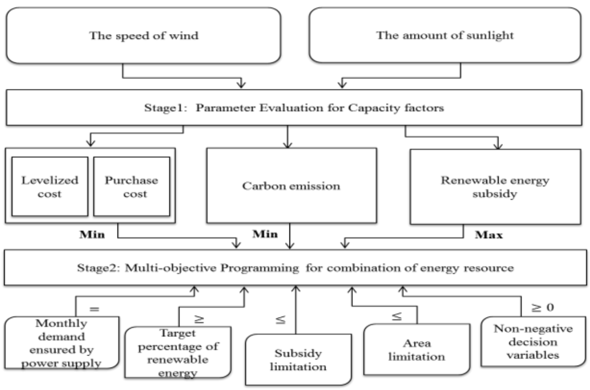


Figure 1 *Framework of Multi-objective Programming with Parameter Evaluation Model*

3.3 Proposed Multi-Objective Programming with Parameter Evaluation Model

Notations on the multi-objective programming with parameter evaluation model are described in section 3.3.1, whereas its objectives and constraints are described in sections 3.3.2 and 3.3.3, respectively.

3.3.1 Notations

Indices

$k \in \{S, W\}$: Solar energy(S), wind energy (W).

$u \in \{S, W, T\}$: Solar energy, wind energy, Taipower (T).

$t \in T = \{1, \ldots, 12\}$: Monthly time interval.

Decision Variables

x^k: Installed capacity of renewable energies. (KW)

y_t: Purchased energy from Taipower in month t. (kWh)

Parameters

$time^t$: Total operation hours in month t.

CF_t^k: Capacity factor of renewable energies in month t.

LC^k: Levelized cost of unit electricity generation of renewable energies. ($\$/kWh$)

CAP_t: Capital expenditures of generation system in month t. ($\$/\mathrm{KW}$)

OPE_t: Fixed operations and maintenance expenditures of generation system in dollars per kilowatt in month t. ($\$/\mathrm{KW}$)

FUE_t: Variable fuel expenditures of non-renewable energy in dollars per kilowatt in month t. ($\$/\mathrm{KW}$)

r_t: Discount rate in month t.

n Life of the system.

C^T: Purchased price from Taipower per unit of power generation

CV^u: Conversion factor of carbon emission. (kg/kWh)

EC^u: External cost of carbon emission. (kg/kWh)

XC^u: External costs of pollutants of solar energy. (kg/kWh)

P^k Equipment subsidy of renewable energies by Taipower

U^k Total subsidy limitation of renewable energies by Taipower

D^t: Electricity demand in month t. (kWh)

K: Total available area can be used

Q^k: Target percentage of renewable energy

v_C: Cut-in speed of wind turbine

v_R: Rated speed of wind turbine

v_F: Cut-off speed of wind turbine

v: Monthly average speed

G_{opt}: Solar irradiance

PR: Performance ratio of solar PV plant modules

$P_{max,STC}$: Total installed capacity of solar energy

H_t: Total generation hours of solar PV plant modules in a time period

3.3.2 Parameter Evaluation with Probability Distribution Function

1. Evaluation Model for the Capacity Factor of Solar Energy

The equation for the capacity factor is shown as Eq (1) (Milosavljević et al., 2015). From the equation, the actual amount of electricity is equal to the product of the amount of solar irradiance and performance ratio. The value of power output can be estimated using installed capacity and operation hours.

$$CF^{S} = \frac{Y_F}{H_t} = \frac{E_{opt}}{P_{max,STC}.H_t} = \frac{G_{opt}.PR}{P_{max,STC}.H_t} \tag{1}$$

where E_{opt} is the power generation of solar and transmitted to the power grid; $P_{max,STC}$ is total installed capacity of solar energy; G_{opt} represents solar irradiance; and H_t is total generation hours in a time period.

2. Evaluation Model for the Capacity Factor of Wind Energy

The function of the Weibull distribution contains two parameters. One is the scale parameter c and the other one is the shape parameter k.

$$f(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} e^{-\left(\frac{v}{c}\right)^{k}}, \quad v > 0, k > 0, c > 1 \tag{2}$$

Let $x = \left(\frac{v}{c}\right)^{k}$, $dx = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1} dv$, $v = cx^{\frac{1}{k}}$, and take into account Eq. (2). The function will be transformed into

$$CF^{W} = \frac{c^3}{v_R^3} \int_{\left(\frac{v_C}{c}\right)^k}^{\left(\frac{v_R}{c}\right)^k} x^{\frac{3}{k}} e^{-x} dx + \int_{\left(\frac{v_R}{c}\right)^k}^{\left(\frac{v_F}{c}\right)^k} e^{-x} dx \tag{3}$$

where v_C is the cut-in speed; v_R is the rated speed; and v_F is the cut-off speed.

3.3.3 Multi-Objective Model

The objectives of renewable energy combination from the aspects of economy, environment, and sustainability are defined in this subsection. Consequently, we adopt a multi-objective optimization that includes the goals of cost minimization, carbon reduction, and sustainable power supply to determine a solution that considers the three aspects mentioned above.

1. Economic Aspect: Eq. (4) can be used to evaluate the annual generation cost of renewable energy and purchase cost from the power company to achieve economic optimization.

$$\text{Min} \sum_{k=S,W} \sum_{t=1}^{12} \left[\left(x^k \times CF_t^k \times time^t \right) \times LC^k \right] + \sum_{t=1}^{12} \left(y_t^T \times C^T \right) \tag{4}$$

where LC^S generation is cost of solar power; LC^W represents the generation cost of wind power; and C^T is the purchase price from the power company. For renewable energy, we considered the costs of construction as well as the operation and maintenance costs without the cost of fuel. We can then multiply the hourly power generation by the total operation hours and leveled cost of electricity of renewable energy to obtain the total generation cost. As to the power bought from the power company, we merely calculated the total purchase expenditure.

2. Environmental Aspect: We usually use the cost of carbon emissions to measure the level of environmental impact. However, the environmental costs of power generation system include generator manufacturing, power generation, and waste disposal. The life cycle assessment is a common approach to accumulate and evaluate data on the inputs, outputs, and the potential environmental effect from the production of raw materials to the final disposal. The environment cost is shown in Eq. (5).

3. Sustainability Aspect

$$\min \sum_{k=S,W} \sum_{t=1}^{12} \left[\left(x^k \times CF_t^k \times time^t \right) \times \left(CV^k \times EC^k + XC^k \right) \right] + \sum_{t=1}^{12} \left[y_t^T \times \left(CV^T \times EC^T + XC^T \right) \right] \tag{5}$$

where *CV (kg/kWh)* represents the conversion coefficient by per unit of power generation of carbon emission for each energy resource; *EC ($/kg)* is the external costs of carbon emission to be charged per kilogram; and *XC ($/kWh)* is the external cost of pollutants for each energy resource.

4. Sustainability Aspect: Because of the self-production of renewable energy, maximizing the amount of renewable electricity is the result we want to reach when the supply and demand is balanced. Hence, government subsidies are then incorporated into the objective function to expand renewable energy generation and to provide a sustainable power supply system.

$$Max \sum_{k=S,W} \sum_{t=1}^{12} (x^k \times P^k) \tag{6}$$

where P^S is the equipment subsidy of solar energy by Taipower and P^W is the equipment subsidy of wind energy by Taipower.

3.3.4 Constraints and Restrictions

1. Monthly power demand should be satisfied by monthly power supply

$$x^S \times CF_t^S \times time^t + x^W \times CF_t^W \times time^t + y_t^T = D_t, \forall t = 1,....,12 \tag{7}$$

where D^t denotes electricity demand in month t.

2. The maximum available area can be used to construct a renewable energy generation system.

$$\frac{x^S}{A^S}+\frac{x^W}{A^W}\le K \tag{8}$$

3. The target percentage of renewable energy should be based on government policy.

$$x^S \ge Q^s * \sum_{t=1}^{12} y_t^{\ T} \tag{9}$$

$$x^W \ge Q^W * \sum_{t=1}^{12} y_t^{\ T} \tag{10}$$

4. The total subsidy limitation of renewable energy should be established according to government policy.

$$x^S \times P^S \le U^S \tag{11}$$

$$x^W \times P^W \le U^w \tag{12}$$

5. Non-negative constraints should be established for monthly electricity purchased from the power company as well as for the installed capacity of renewable energies.

$$x^S \ge 0 \tag{13}$$

$$x^W \ge 0 \tag{14}$$

$$y_t^T \ge 0 \tag{15}$$

3.4 Discussion and Conclusion

We propose a two-stage analytical model to deal with the renewable energy combination. At the first stage, the parameter evaluation model is developed to determine the capacity factor of renewable energies for each month. At the second stage, a multi-objective model is proposed to optimize the mix of electricity from the power utility company and renewable energy generated by the private enterprise. The multi-objective model has three objectives. The generation cost is measured in the economic aspect; the environmental aspect is addressed through the carbon emission cost; whereas and the sustainability aspect is addressed by government policies seeking the balance between supply and demand to ensure security.

4. Numerical Illustration

4.1 Background Description

The National Tsing Hua University (NTHU) of Taiwan is taken as our case study. Taiwan has only one power utility, the Taiwan Power Company (Taipower). Hence, residents, industries, and government agencies have to buy electricity from Taipower.

The Taiwanese government has begun to develop renewable energy generation systems and has formulated subsidy policies to encourage private generation of renewable energies.

Additionally, the Executive Yuan has encouraged government agencies and schools to carry out carbon reduction plans. The Executive Yuan has initiated a program for saving power, fuel, water and paper to increase energy conservation, and expect government agencies and schools provide guidance to civil society in terms of adopting energy saving measures (Bureau of Energy, Ministry of Economic Affairs, 2014).

With the rising awareness on the importance of energy saving, NTHU has also considered establishing its own renewable energy power plant and contribute to the country's efforts at environmental protection. Because a large-scale wind turbine poses dangers and noise problems, a small wind turbine will be used in our case study. Roof type solar panels are also more appropriate for use in the campus to develop renewable energy generation system. The available roof area of NTHU is 79.61×10^3 square meters, which will serve as our land resources for the renewable power plant and to implement comprehensive carbon reduction. The proposed optimal combination renewable energy system was applied, and the Gurobi 5.6.3 software is used for the analysis.

The following sections introduce the input data, including monthly average temperature, monthly global irradiance, monthly average wind speed, cost data, and demand data of NTHU.

1. Solar Irradiance

Table 3 *Monthly Averaged Solar Irradiance of NTHU*
(Source: NASA Atmospheric Science Data Center, 2015)

Monthly average insolation incident on a horizontal surface (kWh/ m^2 /day)					
Jan.	Feb.	Mar.	Apr.	May.	Jun.
2.63	3.01	3.53	4.26	4.88	5.70
Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
6.69	6.11	5.11	4.12	3.10	2.67

2. Wind Speed

Table 4 *Monthly Averaged Wind Speed of NTHU*
(Source: NASA Atmospheric Science Data Center, 2015)

Monthly Average Wind Speed at the Surface Of The Earth (m/s)					
Jan.	Feb.	Mar.	Apr.	May.	Jun.
8.57	7.97	6.76	5.80	5.03	5.39
Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
4.88	5.01	6.30	8.14	8.88	8.52

3. Conversion Factor

Table 5 *Greenhouse Gas Emission of Various Power Generation Technologies (Source: Intergovernmental Panel on Climate Change, 2014)*

Greenhouse Gas Emission within life cycle analysis assessment (g- CO_2 eq/kWh)			
	Low	Average	High
Coal-PC	740	820	910
Gas-combined cycle	410	490	650
Biomass-cofiring	620	740	890
Biomass-dedicated	130	230	420
Geothermal	6.0	38	79
Hydropower	1.0	24	2200
Nuclear	3.7	12	110
Concentrated solar power	8.8	27	63
Solar PV-rooftop	26	41	60
Solar PV-utility	18	48	180
Wind onshore	7.0	11	56
Wind offshore	8.0	12	35

4. Carbon Tax

The carbon tax of 0.54($NT/kg) as imposed by South Korea is used in our study (IEA, 2011, World Energy Outlook 2011).

5. Cost

Table 6 *Levelized Cost of Renewable Energies in Year 2016 (Source: Industrial Technology Research Institute, 2014, Taiwan 2050 Calculator)*

Levelized generation cost ($NT/kWh)					
Onshore Wind	Onshore Wind	Offshore wind (Small)	Hydro	Solar (Roof)	Solar (Ground)
2.83	6.25	8.00	1.24	7.22	6.13

6. Land use

Table 7 *Required Installation Land of Renewable Energies*

Required Installation Land (m^2/KW)		
Onshore wind (400W)**	Onshore wind (3KW)*	Solar (roof)***
3.25	1.40	10
* **Estimated source: TECO ECO Business Group, 2016;*** Source: Industrial Technology Research Institute, 2016		

7. Demand

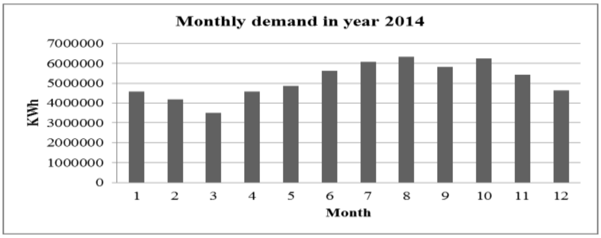


Figure 2 *Monthly Electricity Demand in Year 2014 (Source: National Tsing Hua University, 2014)*

4.2 Illustration of Proposed Parameter Evaluation Model

4.2.1 Parameter Evaluation of Solar Energy

We use Eq. (1) and solar irradiance data from Table 3 to obtain the capacity factor of solar energy at each month. Figure 3 illustrates the relationship between capacity factor and solar irradiance.

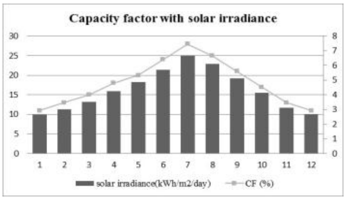


Figure 3 *Monthly Capacity Factor of Solar Energy at Different Solar Irradiance*

Figure 3 shows that the correlation of capacity factor and solar irradiance is highly positive. The variations in capacity factor can be attributed to the performance ratio, which is determined by PV system and reference yields.

4.2.2 Parameter Evaluation of Wind Energy

We use the monthly averaged wind speed at 50m above the surface of the earth, which was obtained from NASA Atmospheric Science Data Center, to determine the probability distribution diagram. The two main parameters are c=6.78 and k=3.00 as shown in Figure 2.

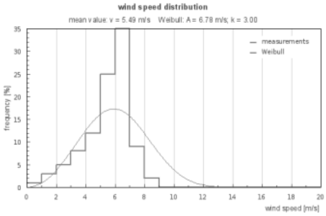


Figure 4 *Weibull Distribution Diagram of Wind Speed*

A higher cut-in speed is necessary because of the smaller capacity of wind power. Hence, we used Aventa AV-7, which has a cut-in speed of 1m/s. The three parameters of Aventa AV-7 are cut-in speed *(V_C=1m/s)*, rated speed *(V_R=5m/s)*, and cut-off speed *(V_F=15m/s)* (Swiss Wind Power Data Website, 2016).

We use Eq. (3) and data from Table 4 to obtain the estimated values of capacity factor of wind energy for each month. The relationship between capacity factor and wind speed is illustrated in Figure 5.

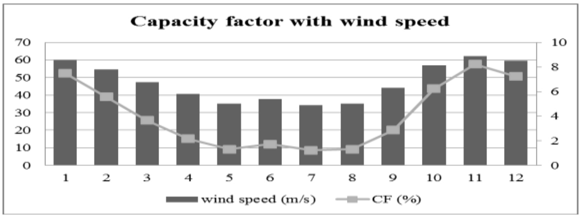


Figure 5 *Monthly Capacity Factor of Wind Energy at Different Wind Speed*

As mentioned in the literature review, the capacity factor is proportional to the cube root mean of the wind speed. From Figure 5, we can see that the capacity factor of wind energy is very low with respect to low wind speed. However, the capacity factor of wind energy rapidly increased at high wind speed.

4.2.3 Conclusion of Parameter Evaluation Model

We provide a numerical example to illustrate the results of our parameter evaluation model. The example shows that the trend of capacity factor of wind is lower in the summer and higher in the winter. On the contrary, the capacity factor of solar energy presents seasonal variations, especially from the June to October, which is the highest within the entire year. These two types of renewable energy can provide a promising compliment to electricity generation.

Based on the appropriate probability density distribution of wind speed with k=3 and c=6.8, which are adopted to evaluate the capacity factor of wind power, and the performance ratio with solar irradiance to estimate capacity factor of solar energy, we can observe that the relationship between the estimated value and the parameters is consistent with the evaluation method. From solar energy, the correlation of capacity factor and solar irradiance is highly positive. Whereas from wind energy, the capacity factor is very low when wind speed is also lower, but increases rapidly at high wind speed.

4.3 Illustration of Multi-Objectives Model

4.3.1 Structure of Goal Programming

The NTHU has begun to develop a renewable energy program in response to the current policies of the Taiwanese government. Thus, we take the university as an example. The goal programming model can be used to meet the target percentage of renewable energy as near as possible. As a model that seeks to comply with government policy, we strive to achieve the goal of environmental protection, and then to look for minimum costs. The three goals are preemptively ordered as follows: environment > economy > sustainability, in which ">" denotes importance precedence. The target values and deviation variables are shown in Table 8.

Table 8 *Goals and Deviation Variables under Government Policy*

Preemptive	Target Value (NT$/year)	Deviation Decision Variable
1. Environment	$Z_2 \leq 55.93 \times 10^6$	d_2^+
2. Economy	$Z_1 \leq 20.00 \times 10^7$	d_1^+
3. Sustainability	$Z_3 \geq 25.00 \times 10^5$	d_3^-

Since the new buildings were opened for use, the electricity expense of NTHU has reached over $19.00 \text{x} 10^1$ million. Hence, we hope that part of the electricity supplied by renewable energy sources can reduce the surcharge of demand in 2014 (National Tsing Hua University, 2014). The target value of environment objective should be below 200 million.

The environmental cost is proportional to the generation structure of the energy sources, which corresponds to $42.96 \text{x} 10^{-2}$% of solar energy and $76.40 \text{x} 10^{-2}$% of wind power in 2016 (Industrial Technology Research Institute, 2014, Taiwan 2050

Calculator). If we want to achieve the target percentage of the government, then the target value of sustainability should be more than 25.00×10^5 NTD per year based on the maximum subsidy of 50.00×10^6 NTD of solar PV within its lifetime.

Suppose that habits of consuming electricity are the same in 2014 and 2015. The price of electricity which will be applied is according to electricity expense (2.97 NTD/kWh) of NTHU in 2015. Because the remaining electricity we should purchase from power company is almost same as NTHU in 2015.

4.3.2 Output of Renewable Energy Combination

The results of goal programming are evaluated in sequence.

- First preemptive order: Minimize the amount of external cost deviation from target value.

Table 9 *Renewable Energy Combination of 1st Preemptive Order*

Renewable Energies	Install Capacity
Solar Energy	82.30×10^1 KW
Wind Energy	46.98×10^1 KW

The deviation of external cost is 0, which means that the first preemptive order is achieved. Therefore, we can continue to evaluate the second preemptive order from the economic aspect.

- Second preemptive order: Minimize the amount of generation cost deviation from target value.

Table 10 *Renewable Energy Combination of 2nd Preemptive Order*

Renewable Energies	Install Capacity
Solar Energy	81.82×10^1 KW
Wind Energy	61.28×10^1 KW

The deviation of generation cost is 0, which means that the second preemptive order is achieved. Therefore, we can continue to evaluate the third preemptive order from the sustainable aspect.

- Third preemptive order: Maximize the amount of subsidy deviation from target value.

Table 11 *Renewable Energy Combination of* 3rd *Preemptive Order*

Renewable Energies	Install Capacity
Solar Energy	10.00×10^2 KW
Wind Energy	51.51×10^1 KW

The deviation of subsidy is 0, which means that third preemptive order is achieved. Therefore, the goal programming model is finished with the complete achievement of the three objectives. Although the external cost of solar energy is higher than wind energy, the generation cost of wind energy is much more than that of solar energy. Therefore, the installed capacity favors solar power under the optimal combination.

4.3.3 Renewable Energy Combination Analysis

1. Purchased Electricity from other Utilities

In addition to the amount electricity generated from renewable energy, the monthly purchased electricity from Taipower also should be considered to reach a balance between supply and demand.

Figure 6 shows the complementarity between purchased electricity and wind power generation in the winter because of the high efficiency of wind power generation. The months of March to August represented not only the peak demand for electricity, but also the peak of the intermittence of wind power. Monthly capacity factors of solar energy are not high for the entire year. Therefore, purchased electricity continued to increase with the demand. Although the installed capacity of solar energy is much higher than that of the wind energy, the high capacity factor of wind energy generated higher power output for NTHU in winter.

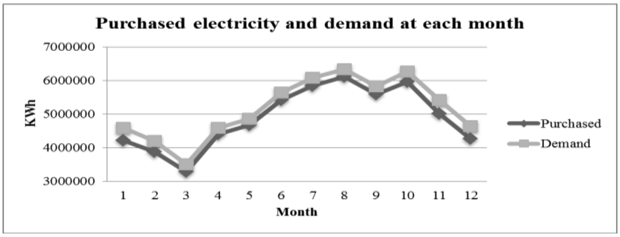


Figure 6 *Monthly Purchased Electricity and Demand in Year 2015*

2. Output Values of Objectives

After deciding on the optimal combination of renewable energy, we can investigate the corresponding value of each energy source.

Table 12 *Objective Values*

Objective Values	Total Cost (NT$/year)
Generation cost of solar	11.40×10^6
Generation cost of wind	10.29×10^6
Purchase cost from Taipower	16.83×10^7
Total internal costs	19.00×10^7
External cost of solar	59.55×10^4
External cost of wind	60.40×10^3
External cost for Taipower	53.86×10^6
Total external costs	54.52×10^6
Equipment subsidy of solar	25.00×10^5
Equipment subsidy of wind	0
Total subsidies	25.00×10^5

The cost comes mostly from electricity production, of which per unit renewable energy is much higher than the price set by the power company. The total generation cost of solar energy is nearly the same as that of wind energy despite the optimal combination of installed capacity of wind being approximately $50.00\text{x}10^1$ kilowatts less than solar energy. These numbers indicate that the original generation cost of wind power is very high. The external cost is composed of both carbon tax and hazard cost as shown in Table 1. The per unit external cost of solar energy is nine times higher than that of wind energy, but the installed capacity is about twice that of wind energy. From the objective values, we can observe that the external cost is an important consideration because each energy resource has a different carbon emission coefficient. If we multiplied the total amount external costs of the emissions with the carbon emission coefficient, the environment objective will be caused by larger increments of cost. Table 5 shows that the emission coefficient of coal is 70 times more than that of onshore wind with respect to the mean value. The power generation structure of Taipower can be decomposed as follows: coal-fired (37.60%), fuel oil (28.00x10^{-1}%), fuel gas (32.40%), co-generation (32.00x10^{-1}%), renewable (54.00x10^{-1}%), and nuclear (18.60%). Therefore, the emission coefficient of this generation system is equal to 47.80x10^{-2} (kg/kWh) claimed by Taipower. Compared to renewable energy, that for solar and wind energies are both 0.041 (kg/kWh) and 0.011 (kg/ kWh), respectively. Consequently, carbon emission is the most sensitive factor that we should seriously consider if we decide to proceed with renewable energy penetration in our study.

Table 13 shows the total electricity consumption cost and total external cost investigated of NTHU in 2015.

Table 13 *Electricity Cost Structure of NTHU. (NTHU, 2015)*

Objective Value	Total Cost (NT$/year)
Purchase cost from Taipower	17.65×10^{7}
External cost for Taipower	56.48×10^{6}

A comparison of Tables 12 and 13 shows that the internal cost after developing renewable energy generation system will reduce by NT$ 26.21x10^{5} and that the external costs will be also reduced by NT$ 14.06x10^{5} in one year. Consequently, the reduced internal cost becomes higher than the reduced external cost when considering carbon emissions. Because the price of electricity from Taipower is charged according to different levels of electricity consumption, an additional fee is imposed when the total consumption expenditure from Taipower is exceeded. Hence, enterprises with high demand for electricity can obtain more benefits if it uses a renewable energy generation system.

4.4 Sensitivity Analysis

We perform a sensitivity analysis to discuss the effects of the optimal solution when the coefficient changes. From the output in section 4.3, we can see that the external cost of energy source remains a prominent portion of total cost because the environment hazards of power generation are inevitable. However, because renewable energy is being promoted in a large number of countries worldwide, carbon tax will also be seriously considered worldwide.

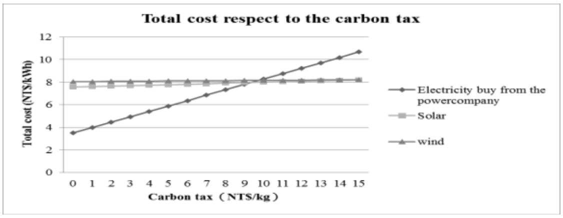


Figure 7 *Total Cost Increment respect to Carbon Tax*

With the increase in carbon tax, the total cost of renewable energy will also slowly rise, whereas purchasing cost and external costs of fossil fuels from Taipower will increase rapidly. Initially, the total cost of purchasing electricity from Taipower will be less than that of solar energy, but this cost will result in an increase in carbon tax. The reason is that the power structure of Taipower comes mostly from fossil fuels, and the carbon emission coefficient of fossil fuels is much higher than that for renewable energy. Hence, we should investigate how these costs affect the optimal solution.

Table 14 *Optimality Interval of Carbon Tax*

Interval of Carbon Tax	$0 \le t < 97.40\times10^{-1}$		$97.40\times10^{-1} \le t < 15.67$		$15.67 \le t < \infty$	
Installed Capacity(Kw)						
$X = (x^S, x^W)$	x^S	10.00×10^2	x^S	78.89×10^2	x^S	10.00×10^2
	x^W	51.51×10^1	x^W	51.51×10^1	x^W	49.72×10^3
Objective value						
$Z = (Z_1, Z_2, Z_3)$	Z_1	19.00×10^7	Z_1	23.75×10^7	Z_1	10.05×10^8
	Z_2	54.52×10^6	Z_2	31.41×10^7	Z_2	42.90×10^6
	Z_3	25.00×10^5	Z_3	25.00×10^5	Z_3	25.00×10^5

The optimal combination of renewable energies will change when carbon tax reaches beyond $97.40\text{x}10^{-1}$ NT$/kg. The external cost for Taipower will be higher than that of renewable energy because it will cause a necessary increase in solar installation capacity. The increment of external cost is also inevitable because of the increase in carbon taxes. Therefore, the optimal combination of renewable energy will have the lowest cost in the case of NTHU when carbon tax less than $97.40\text{x}10^{-1}$ NT$/kg. If the carbon tax goes beyond $97.40\text{x}10^{-1}$ NT$/kg, both of the values of objectives 1 and 2 will also increase. This increment will result in an increase of NT$ $47.58\text{x}10^{6}$ in generation cost because of the increasing total generation cost of solar energy. The total external cost will decrease to NT$ $42.90\text{x}10^{6}$ when the carbon tax goes beyond 15.67 NT$/kg. The reason for this result is because wind power has the lowest carbon emission for its increment. The equipment subsidy will remain at the same value because it has achieved the upper bound of subsidy. Based on the results and with the prospect of a future carbon tax policy, the development of wind power will be the most beneficial in terms of reaching the carbon reduction targets.

5. Conclusion & Future Work

5.1 Summary and Conclusion

With the depletion of fossil fuels, renewable energy substitution has become a trend in electricity supply sources worldwide. Environment protection has also become an urgent issue that needs to be addressed. Renewable energy is sustainable and clean, and many countries have established subsidy policies to encourage enterprises to develop their own renewable energy generation systems.

Electric power plants have several issues, including carbon emission and sustainable power supply. Consequently, the proposed combination of renewable energies model considers three aspects, including that of economic, environmental, and sustainability aspects. In contrast with the conventional energy planning model that focuses on the mix of non-renewable and renewable energies, this research investigates not only the mix of electricity from power companies and renewable energy generation but also the combination of renewable energies.

The results of the case study of NTHU shows that the generation cost of wind power remains too high, and hence, purchased electricity continues to be a large proportion of the power used by NTHU. Carbon emission cost is the main factor that changes the structure of power generation.

If we wish to promote renewable energy, we have to use a higher efficiency technology for solar power. This promotion will not only result in a substantial decline in its cost, but could also enhance carbon capture technology, thereby reducing the cost of carbon emissions. The Taiwanese government should consider formulating subsidy policies and taxation differentiation strategies to increase the proportion of renewable energy power generation.

5.2 Future Work

This study aims to decide on the amount of infrastructure necessary to be installed for each type of renewable energy by considering their proportions. Accordingly, the possible expansion of the power generation system, taxation differentiation strategies, and subsidy policies must be investigated in future research. From the result of chapter 4, we can observe that carbon emission is a sensitive factor because the generation structure will change with increasing tax. Through the analysis, we made a suggestion that a taxation differentiation strategy could be an effective strategy. Because we have to pay the environment costs per unit of electricity, we can set a rule for carbon tax under different scales of generation. In this way, more focus will be given to the environmental benefit from each unit power of renewable energy sources. As long as we consider the effects on the environment caused by different energies and differential carbon taxes, we can control the total external costs in the generation system.

6. References

1. Arslan, T., Bulut, Y.M., Yavuz A., “Comparative study of numerical methods for determining Weibull parameters for wind energy potential”, Renewable and Sustainable Energy Reviews, 2014, pp. 820-825
2. Bilil, H., Aniba, G., Maaroufi, M., “Multiobjective Optimization of Renewable Energy Penetration Rate in Power Systems”, Energy Procedia, 2014, pp.368-375

3. Bureau of Energy, Ministry of Economic Affairs, “Rising Green Energy Industry Program”, 2014
4. Chen, P.Y., Chen, B.Y., Tsai, P.H., Chen, C.C., “Evaluating the impacts of a carbon tax on imported forest products—evidence from Taiwan”, Forest Policy and Economics, 2015, pp.45-52
5. Cheng, H., Huang, W., Zhou, Q., Cai, J., ”Solving fuzzy multi-objective linear programming problems using deviation degree measures and weighted max–min method”, Applied Mathematical Modelling, 2013, pp. 6855-6869
6. Filippo, S., Fabio, C.,” Monitoring and checking of performance in photovoltaic plants: a tool for design, installation and maintenance of grid-connected systems”, Renewable Energy, 2013, pp. 722-732
7. Ho, Y.F., Chang, C.C., Wei, C.C., Wang, H.L., “Multi-objective programming model for energy conservation and renewable energy structure of a low carbon campus”, Energy and Buildings, 2014, pp.461-468
8. International Energy Agency, “World Energy Outlook 2011”, 2011
9. International Energy Agency, “Key World Energy Statistics 2013”, 2013
10. Islam, M.R., Saidur, R., Rahim, N.A., “Assessment of wind energy potantiality at Kudat and Labuan, “Malaysia using Weibull distribution function”, Energy, 2011, pp. 985-992
11. Kwon, S.D., “Uncertainty analysis of wind energy potential assessment”, Applied Energy, 2010, pp. 856-865
12. Milosavljević D.D., Pavlović T.M., Piršl, D.S., ”Performance analysis of A grid-connected solar PV plant in Niš, republic of Serbia”, Renewable and Sustainable Energy Reviews, 2015, pp. 423-435
13. Moura, P.S., Almeida, A.T. de., “Multi-objective optimization of a mixed renewable system with demand-side management. Renewable and Sustainable Energy Reviews”, Renewable and Sustainable Energy Reviews, 2010, pp.1461-1468
14. Nuclear Energy Agency, “System Effects in Low-carbon Electricity Systems”, Nuclear Energy and Renewables, 2012
15. Park, E., Kwon, S.J., “Solutions for optimizing renewable power generation systems at Kyung-Hee University′s Global Campus, South Korea”, Renewable and Sustainable Energy Reviews, 2016, pp. 439-449
16. Pineda, S., Bock, A., “Renewable-based generation expansion under a green certificate market”, Renewable Energy, 2016, pp. 53-63
17. Ragab, M., Som, A.K., “A preliminary study of wind power potential in Bahrain”, Renewable Energy, 1993, pp. 67-74
18. Ryu, H., Dorjragchaa, S., Kim, Y., Kim, K., “Electricity-generation mix considering energy security and carbon emission mitigation: Case of Korea and Mongolia”, Energy, 2014, pp.1071-1079
19. Sharafi, M., ELMekkawy, T.Y., “Multi-objective optimal design of hybrid renewable energy systems using PSO-simulation based approach”, Renewable Energy, 2014, pp.67-79
20. Shen, J.N., “Unit sizing and economic analysis of an autonomous photovoltaic system”, National Science Council,2000
21. Zhang, H., Li, L., Zhou, D., Zhou, P., “Political connections, government subsidies and firm financial performance: Evidence from renewable energy manufacturing in China”, Renewable Energy, 2014, pp.330-336

22. Bureau of Energy, Ministry of Economic Affairs, 2015 <http://web3.moeaboe.gov.tw/ECW/populace/home/Home.aspx>
23. Environment Protection Administration Executive Yuan, Taiwan, 2015 <http://www.epa.gov.tw/mp.asp?mp=epa>
24. State Grid Energy Research Institute, National Energy administration, 2012 <http://www.nea.gov.cn/2012-02/10/c_131402694.htm>
25. The Swiss Wind Power Data Website, 2016. <http://wind-data.ch/tools/powercalc.php?lng=en>
26. TECO ECO Business Group, 2016 <http://eco.teco.com.tw/Modules/Web/Prod.aspx?PROD_ROLES=2>
27. National Tsing Hua University, 2014 < http://nesh.web.nthu.edu.tw/files/11-1021-9305.php?Lang=zh-tw >

7. Acknowledgment

The authors acknowledge the financial support from Ministry of Science and Technology in Taiwan, ROC with project number MOST 103-2221-E007-057.

About Our Authors

Hsiao-Fan Wang is the Emeritus Distinguished Chair Professor of National Tsing Hua University (NTHU), Taiwan, Republic of China and the Fellow of Chinese Institute of Industrial Engineering. She is also taking charge of the Arts Center of the University. Her research interests are in Multicriteria Decision Making, Fuzzy Set Theory and Green Value Chain Management.

Pei-Wen Chou is a master graduate from the Department of Industrial Engineering at National Tsing Hua University (NTHU), Taiwan, Republic of China. She has been doing research on remanufacturing, optimization, and energy resources in Soft Computing and Fuzzy Optimization Lab, National Tsing Hua University. She is now an engineer of TSMC.

Incentive Model for Operational Level Employee Teams in an Organization

Vijitha Ratnayake
Gayathri Perera
University of Moratuwa, Sri Lanka
(vijithar@uom.lk)

Volume 24, *Number* 2
June 2018, *pp.* 139-152

This paper aims to address a significant gap in the available literature on incentive scheme models for indirect employees based on their performance levels. A comprehensive mathematical model is developed for incentivizing the operational level indirect employees by using stochastic gradient descent algorithm. The developed incentive scheme establishes link between three broad levels of an organizational hierarchy; Strategic level, tactical/departmental level and operational level. The major considerations of the developed incentive scheme model are organizational performance, allocation of incentives between departments, incentive period, team performance, target setting for Key Performance Indicators (KPIs) and correlation of KPIs to the organizational performance.

Keywords: Incentive Scheme, Machine Learning, Motivation, Indirect Employee, Performance Measurement

1. Introduction

According to Cox (2014), Goyal (2014) and McDougall & Radvanovsky (2008), in traditional approach, an organization can be represented in three broad levels such as; Strategic level, Tactical level and Operational level. In addition, the entire workforce in an organization can be broadly categorized into two sections as; direct and indirect employees (Bragg, 2009). Financial incentives often result in significant increase of employee motivation and workforce productivity (Chary, 1995). There are multiple systems available in literature on giving incentives for direct employees (Das, 2013; Chary, 2009; Aswathappa, 2005). However, there is lack of indirect employee incentive schemes due to the difficulty in accurate performance measurement of the indirect employees (Das, 2013; Chary, 2009; Lal & Srivastava, 2009; Alan & Pizzey, 1989).

In this paper, a comprehensive mathematical model is developed for incentivizing the operational level indirect employees. Team based financial incentive scheme is developed by considering both employer and employees' perspectives. Stochastic gradient descent algorithm is used for the mathematical formulation. The developed model ensures a consistent link between abovementioned three levels of an organization, which ultimately creates win-win solutions to both the employees' and employer's parties.

2. Theoretical Background

2.1 Definition of Direct and Indirect Employees

According to Bragg (2009), direct employees are engaged directly in the manufacturing process of final output. The traceable processes, which include changing of the composition, form or condition of the final product is done by direct employees. Conversely, any employee involved in facilitating, maintaining and supervising the total manufacturing process are categorized as indirect. Only the direct employees are considered as engaged in direct value addition processes (Bragg, 2009). However, it is obvious that direct workforce cannot be fully utilized and productive without the contribution of indirect employees (Das, 2013; Lal & Srivastava, 2009; Alan & Pizzey, 1989). The higher performance level of direct employees is achievable only with the effective cooperation of indirect employees (Das, 2013; Rao, 2007). The direct employees will receive smoother flow of production with the positive intervention of the indirect employees (Banerjee, 2006). However, there is a lack of indirect employee incentive schemes due to the difficulty in accurate performance measurement of the indirect employees (Das, 2013; Chary, 2009; Lal & Srivastava, 2009; Alan & Pizzey, 1989). Therefore, the performance based monetary evaluation of indirect employees is essential in achieving the ultimate goals and objectives of the company (Chary, 2009; Rao, 2007).

2.2 A Financial Incentive Scheme: Importance and Constraints

Incentive scheme can be either financial or non-financial. As stated by Grammling & Holtmann (2006), the financial incentive schemes have more impact on the performance of the employees than that of non-financial incentives. Although, financial incentives are often considered as an overburden to the employers, they are leading to significant increase of workforce productivity (Chary, 1997). However, financial incentive schemes can possibly fail if they are not properly designed. The main causes for such failures are; employees crossing the ethical boundaries and disregarding the quality of work in order to achieve incentives, de-motivation due to unrealistic targets & inequality of received incentives even for same job performance and gradual reduction of motivation after achieving required level of performance to gain incentives. (Banfield & Kay, 2012; Perry et al., 2009; Burgess & Ratto, 2003; Kohn, 1993) Therefore, it is essential to introduce an appropriate, well-designed financial incentive scheme to the employees of any organization in order to achieve higher productivity levels.

2.3 Characteristics of a Good Incentive Scheme

Characteristics from Employees' Direction

As per Holtmann (2002), for any incentive scheme to be successful, those who are affected should accept it. Therefore, the designing of an incentive scheme should be done in such a way that will not create an adverse effect on the employees and the company. To become an effective incentive scheme it should fulfil two basic requirements, fairness and transparency (Grammling & Holtmann, 2006). Fairness can be further categorized as; distributive fairness and procedural fairness. The distributive fairness implies that an incentive scheme should allocate the incentives based on the effort paid by the employee. Procedural fairness denotes that an incentive scheme must have a trustworthy, logical and consistent allocation procedure. Incentive calculation procedure must possess clarity and equity. Equity is referred as setting attainable targets within the reach of an average performing employee with noticeable effort. In addition to that, there should be significant difference between the incentives earned by employees with higher performance when compared to low performing employees (Gordon & Kaswin, 2010). Otherwise, the high performing employees will dissatisfy with the feeling of underestimation of their skill levels.

Another key factor, which should be considered in developing an incentive scheme, is the status consistency. The allocated amount of incentives should reflect the employees' position in the organizational hierarchy (Holtmann, 2002). According to the definitions provided by Allen (2011), the operational level consists of staff employees who operate below the executive level. This categorization is done mainly based on the ability to engage in decision-making and controlling power. Common features of operational level are the least independent nature in decision making and routine job roles with minor alterations. However, since the operational level is taken into consideration, which represents a single layer of the organizational hierarchy, the status consistency will not create a significant effect on this model.

According to Holtmann (2002), there is a significant impact on employee productivity with shorter payment intervals. Having shorter time intervals between rewarding will ease the employees when linking the earned incentives to their efforts. Therefore, the corrective actions for poor performance can be taken within short period of time. However, it is undesirable to have very short time intervals due to difficulties in collecting and maintaining the relevant data. Therefore, the incentive scheme for operational level indirect employees in any organization should be having an average time intervals between incentive payments. But, some of the teams may require different incentive calculation period than the others. It happens depending on the nature of their job role. Therefore, the incentive scheme model must have the flexibility of adjusting the incentive calculation period as per need.

Characteristics from Employers' Direction

In organizational employers' point of view, there are two major factors being considered when offering an incentive scheme to the employees. Usually, it is done with the intention of achieving certain level of positive impact on the performance level of the employees whilst maintaining minimum financial expenses (Holtmann & Grammling, 2005). Therefore, managerial level expects an incentive scheme to have the ability of fulfilling above requirements. A well-designed incentive scheme can result in giving several benefits to the employers. The incentive scheme can be considered as a tool to compete with other companies. As a result of a proper incentive scheme, the employees are motivated to achieve higher productivity level and retain within the company while attracting the outsiders with higher capabilities (Holtmann & Grammling, 2005). This will result in enhancing the human resources potential within the company. In addition to that, the company can gain social and customer recognition as a fair pay organization. Such recognition can be used as a marketing tool in both customer segments as well as in local labor market. Therefore, implementation of an employee incentive scheme can render benefits to both employees and employers providing a win-win solution.

2.4 Team Based and Individual Incentive Schemes

The implementation of team based incentive schemes drastically increases the performance of employees than the individual based schemes. As per the analysis conducted by Condly et al. (2003), the increase in performance due to team based incentives is 48% whereas it is 19% with individual based incentive schemes. Interdependency of job roles ‘between the teams’ as well as ‘within the team’ is key factor to be considered when designing an incentive scheme. An individual based incentive scheme can be applicable only if it is possible to distinguish individual performance (Hoffman & Rogelberg, 1998; Holtmann, 2002).

The skill levels of employees vary from lower level to higher level in any organization. In such situation, offering an individual based incentive scheme by setting targets to be matched with high performers may result in dissatisfaction of lower level performance employees and vice versa (Holtmann, 2002). With the presence of an individual based incentive scheme, the employees tend to maximize their own output other than focusing on the goals and objectives of the company.

This is undesirable especially in organizations where the team spirit and cooperation between co-workers is highly encouraged.

In contrast to the individual incentive schemes, team based incentive schemes encourage the intra-team bargaining, collective and mutual learning of employees (Hamilton et al., 2001). Through that, the final product quality will increase simultaneously with personal growth of employees whilst reducing the need of supervision. However, free riding effect is known as the major drawback of team based incentive schemes. It can be mitigated to certain extent by allocation of minimum possible number of employees per team with close supervision (Holtmann, 2002). In addition to that, the peer pressure can influence the free riding effect to a certain extent (Hamilton et al., 2001).

Characteristics of Team based Incentive Schemes

The stability and interdependency of teams play major roles when designing a team incentive scheme (Hoffman & Rogelberg, 1998). Dependent teams’ performance can be directly influenced by the poor performance of other teams. Therefore, the operational relationship between teams must be taken into consideration when setting targets in team incentive scheme. Otherwise, the teams may get discouraged due to evaluation of performance by including uncontrollable and unaccountable effects.

The established teams may be either part-time or full-time. In case of part-time teams, giving incentives for a particular period may be problematic due to unstable nature of the team. Hence, it is advisable to establish full-time teams depending on the desired incentive calculation period (Hoffman & Rogelberg, 1998).

As stated by Hoffmann & Rogelberg (1998), there can be seven major categories of team incentive schemes. They are; Team gain sharing/ profit sharing, Team goal based incentive system, Team discretionary bonus system, Team skill incentive system, Team member skill incentive system, Team member goal based incentive system, Merit incentive systems. Out of these, the Team skill incentive scheme focuses on rewarding the teams with significant performance development during the incentive calculation period. Such team incentive system can be applicable to an organization in combination with proper performance evaluation criteria. Establishing a system to do a quantitative analysis of the team performance is also essential.

2.5 Factors to be Considered when Developing an Incentive Scheme for Indirect Employees

As stated by Das (2013) and Rao (2007), it is essential to consider the following factors when designing an incentive scheme for indirect employees. They are

- The scheme should achieve all-round efficiency in the organization
- It should be organized to relate the reward with the efforts
- It should be so organized as to encourage maximum efficiency
- The scheme should be guaranteed for a specific period
- It should ensure payment of incentives at regular intervals

2.6 Performance Evaluation Method for an Incentive Scheme

According to Ratnayake et al. (2014), the stochastic gradient descent algorithm can be effectively applicable in evaluating the performance contribution of employees at different levels of an organizational hierarchy. The approach used in their study is used in this paper to evaluate the operational level indirect employees' performance and allocate the performance based incentives to the employees.

3. Methodology

Deductive approach is used for the development of the mathematical model. The theoretical background for the development is based on the findings from available literature.

3.1 Decision on Target Party for Incentives

The review of literature has shown that the suitability of targeting employee teams in calculating incentives rather than individuals. Therefore, the incentive scheme model is developed as a team wise incentive scheme. When deciding the employee teams, homogeneity of the job roles within a team is essential. Therefore, when deciding the incentive allocating teams, it is vital to categorize the operational level indirect employees with same job role in to a same job category.

The Key Performance Indicators (KPIs) for each operational level employee team should be established by using an appropriate performance measurement system. Since the selection of KPIs is dependent on nature of the organization, customizable model is developed so that the user can including respective KPIs as per the requirement. For the ease of representation of the developed model, it is assumed that two KPIs are defined to each operational level employee team.

Development of the incentive scheme was done by ensuring factors stated in section 2.5.

3.2 Consideration on the Performance Contribution along Organizational Structure

In determining incentives for operational level indirect employees, it is important to pay attention on the gain can be obtained to the employer simultaneously. Therefore, it is required to link the operational level employee incentive scheme with the overall company performance. As an example, from employers' direction, it is worthwhile to pay incentives for employees only if overall growth of the profits or budgeted standard hours for the period has been achieved. Therefore, it has been decided to incorporate a measure on corporate level (Strategic level) company performance to the developed incentive scheme. As in Figure 1, department level (Tactical level) has been considered as the next performance evaluation stage.

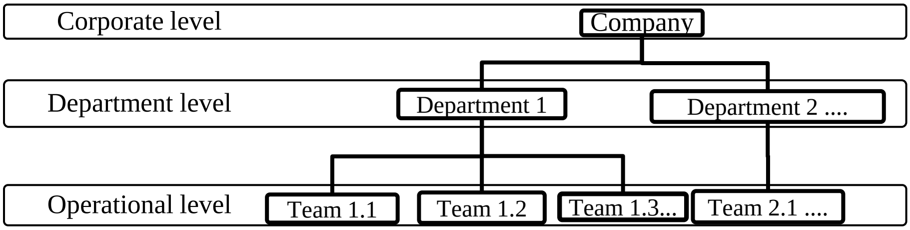


Figure 1 *Performance Contribution Flow through Organizational Hierarchy*

Allocation of the total incentive for organization between departments has to be logical and fair. Allocated incentives for a department have to represent the head count of the department related with the incentive scheme and current period performance of the department. It is important to identify each department's performance contribution for the company's performance of the same period. Team-level performance evaluation and appraisal is the most important part in development of the incentive calculation model. In order to incentive allocation model to be practical, incentives for a team has to show direct relationship with team performance. It is because employees' direct controllability over the incentives is at the stage of team level. Therefore it has been identified that higher the employee head count of the company, more weigh of incentive allocation has to be given at team level.

Accordingly, a team wise incentive scheme has been developed to represent the performance contribution of different levels of a company whilst ensuring that operational level indirect employees are having high controllability over their incentives.

4. Results

4.1 Incentive Scheme Model for an Operational Level Indirect Team Member

Assumptions

1. Departmental KPIs used for the incentive scheme are related to the operational level indirect employee performance.
2. Team level KPIs are directly accountable and controllable to the employee team being incentivized

Customizable incentive scheme model for operational level indirect employees is given in equation (1).

Indexes

i Index for departments ($i = 1,2, \dots, n$)

c Index for operational level indirect employee teams within a department ($c = 1,2, \dots, C$)

α Index for organizational sub-KPIs ($\alpha = 1,2, \dots, A$)

s Index for departmental KPIs (s$= 1,2, \dots, S$)

m Index for KPIs of operational level indirect employee team ($m = 1,2,\dots,M$)

Inputs

A Total number of organizational KPIs

S Total number of departmental KPIs

M Total number of KPIs for operational level indirect employee team

n T Otal number of departments considered under incentive scheme

C Total number of operational level indirect employee teams within a department

KPI_{cm} Value for the KPI m of operational level indirect employee team c during period t

N_i Operational level indirect employee head count in department i belong to incentive scheme

t Adjustment for the incentive calculation period/ days, weeks for the current month

X Total incentive ceiling for organization per incentive calculation period

$KPI_{F,Bud.}$ Budget/ target value of organizational main KPI for the selected period

$KPI_{cm,Bud.}$ Budget/ target values for the KPI m of team c for the selected period

Variables

θ_a Performance contribution weight of organizational KPI α

$\emptyset_{is}$ Performance contribution weight of departmental KPI s of department i

β_{cm} Performance contribution weight of operational level indirect employee team KPI m of team c

$KPI_{F,Act.}$ Actual value of organizational main KPI for the selected period

$KPI_{cm,Act.}$ Actual values for the KPI m of team c for the selected period

KPI_a Weighted value for organizational sub-KPI α during period t

KPI_{is} Weighted value for the KPI s of department i during period t

I_{ic} Total incentive earning of an operational level indirect employee in team c of department i during period t

$$I_{ic} = \frac{\left(\left[X\left(\frac{KPI_{F,Act.}}{KPI_{F,Bud.}}\right)\left(\frac{\emptyset_{is}}{\sum_{i=1}^{n}\emptyset_{is}}\right)\right] + bal.\right)\left(\frac{1}{\sum_{i=1}^{n}N_i}\right)}{t}\left[\sum_{c=1}^{C}\sum_{m=1}^{M}\left(\frac{\beta_{cm}}{\sum_{c=1}^{C}\sum_{m=1}^{M}\beta_{cm}}\right)\left(\frac{KPI_{cm,Act.}}{KPI_{cm,Bud.}}\right)\right] \quad (1)$$

$$KPI_{F,Act.} = \sum_{a=1}^{A}\theta_a KPI_a \quad (2)$$

$$KPI_a = \sum_{i=1}^{n}\sum_{s=1}^{S}\emptyset_{is}KPI_{is} \quad (3)$$

$$KPI_{is} = \sum_{m=1}^{M}\sum_{c=1}^{C}\beta_{cm}KPI_{cm,Act.} \quad (4)$$

$$\left(\frac{KPI_{F,Act.}}{KPI_{F,Bud.}}\right) \leq 1 \quad (5)$$

$$\left(\frac{KPI_{cm,Act.}}{KPI_{cm,Bud.}}\right) \leq 1 \quad (6)$$

As an example, consider an operational level indirect employee team (Team 1) in department 1 having two KPIs, KPI_{11} and KPI_{12}. Assume there is only one organizational level KPI, KPI_1 (where $a = 1$). According to the developed model, the incentive earning per employee of Team 1 - Department 1 for the period t is I_{11} which is given in equation (7).

$$I_{11} = \frac{\left(\left[X\left(\frac{KPI_{F,Act.}}{KPI_{F,Bud.}}\right)\left(\frac{\emptyset_{11}}{\sum_{i=1}^{n}\emptyset_{is}}\right)\right] + bal.\right)\left(\frac{1}{\sum_{i=1}^{n}N_i}\right)}{t}\left[\left(\frac{\beta_{11}}{\sum_{c=1}^{2}\sum_{m=1}^{2}\beta_{11}}\right)\left(\frac{KPI_{11,Act.}}{KPI_{11,Bud.}}\right) + \left(\frac{\beta_{12}}{\sum_{c=1}^{2}\sum_{m=1}^{2}\beta_{12}}\right)\left(\frac{KPI_{12,Act.}}{KPI_{12,Bud.}}\right)\right] \quad (7)$$

Assumption: KPI_1, $KPI_{11,Act.}$ and $KPI_{12,Act.}$ are positively co-related with the company performance. For negatively co-related KPIs, respective adjustment has to be made as given under the section 4.8.

Above incentive earning (given in equation 7) for an operational level indirect employee working at a team of department 1 can be further elaborated through step by step calculations as given under the sections 4.2 to 4.8.

4.2 Adjustment on Organization Performance

The proposed incentive model suggests in maintaining a constant amount of total incentive ceiling (X) per month. It is the maximum amount company would have to pay for employees under the scheme given that all the corporate level performance indicators, department level performance indicators and team level performance indicators have achieved their targets for the month.

If there are multiple sub-KPIs available in corporate level, it is possible to calculate the value for organization's main KPI by using stochastic gradient algorithm given in equation (2).

Use of a constant total incentive ceiling has given the employees a guaranteed incentive distribution if company performs well for the period. Therefore from the employees' perspective, it is encouraging to have an incentive scheme linked with company overall performance, measured through company's profitability or customer-paid standard hours.

Maximum total incentive amount allocated to the organization for current month (X');

$$X' = X\left(\frac{KPI_{F,Act.}}{KPI_{F,Bud.}}\right) \quad (8)$$

From the employer's perspective, it is beneficial because total employee incentive allocation fluctuates in line with the realized return from the organization.

4.3 Allocation of Incentives between Departments

Two major concerns have been taken in to account in allocating the total incentives between departments. They are; employee head count related with the incentive model and departmental performance contribution towards organizational performance for the period.

$$KPI_a = \sum_{i=1}^{n} \sum_{s=1}^{S} \emptyset_{is} KPI_{is} \tag{3}$$

Each department's performance contribution for the organization performance can be found using a regression approach. With a stochastic gradient descent algorithm, above relationship can be solved using training examples (historic data of KPIs). With the values for$\emptyset_{is}$, each department's performance contribution can be found as a ratio (e.g.: for department 1 = $\frac{\emptyset_{11}}{\sum_{i=1}^{n} \emptyset_{is}}$). More localized optimal solution can be attained under this method which gives accurate results to the current incentive calculation period's performance contribution.

However, it is required to pay more weight towards the employee head count at this stage than the departmental performance contribution. It is due to the fact that more priority of performance evaluation of employees should be given to the team level which is under their direct controllability. Therefore in allocating the calculated X' value between departments, relationship has to be developed to give more priority towards employee head count of the departments and less priority towards departmental performance contribution towards corporate level. Also the relationship should consider the fact that higher the employee head count of organization, lower will be the controllability of a departmental performance to a single team. It requires increasing the priority towards employee head count over performance contribution in allocating incentives to departments, when total organization employee head count increases.

Total maximum incentive allocated to department 1 employee for the current month

$$X" = X' \left(\frac{\emptyset_{11}}{\sum_{i=1}^{n} \emptyset_{is}} \right) \left(\frac{N_1}{\sum_{i=1}^{n} N_i} \right) + \ bal. \left(\frac{N_1}{\sum_{i=1}^{n} N_i} \right) \tag{9}$$

Where;

$$bal. = \ X' - \ X'^{\frac{\sum_{i=1}^{n} \emptyset_{is} N_i}{\left(\sum_{i=1}^{n} \emptyset_{is}\right)\left(\sum_{i=1}^{n} N\right)}} \tag{10}$$

In order to avoid unhealthy competition between teams within a same department, equal incentive opportunity has to be given to all teams within a department.

Hence, Maximum current month personal incentive ceiling for an operational level employee of department

$$1\ (M) = \frac{X''}{N_1} \tag{11}$$

$$M = \frac{X' \left(\frac{\emptyset_{11}}{\sum_{i=1}^{n} \emptyset_{is}} \right) \left(\frac{N_1}{\sum_{i=1}^{n} N_i} \right) + bal. \left(\frac{N_1}{\sum_{i=1}^{n} N_i} \right)}{N_1} \tag{12}$$

$$M = \left(\left[X \ \left(\frac{KPI_{F,Act.}}{KPI_{F,Bud.}} \right) \left(\frac{\emptyset_{11}}{\sum_{i=1}^{n} \emptyset_{is}} \right) \right] + \ bal. \right) \left(\frac{1}{\sum_{i=1}^{n} N_i} \right) \tag{13}$$

4.4 Adjustment to Period

An operational level employee team within a department may require in incentivizing daily, weekly or monthly. Therefore considering the number of working days for current month or working weeks for current month, maximum personal incentive for the period can be calculated as below.

Current period personal incentive ceiling for an employee of department 1 (M');
$M' = \frac{M}{t}$

Here t is the adjustment for incentive calculation period. As an example if the KPI values are given monthly and the incentive calculation is required in monthly/weekly/daily basis, the value for t is; t =1 for monthly incentive ceiling, t = 4 for weekly incentive ceiling, t = 20 for daily incentive ceiling (assuming number of working days per month is 20).

4.5 Adjustment on Team Performance

As aforementioned, it is assumed that two KPIs are defined for each operational level employee team. The number of KPIs can be selected as the user requirement.
Two main considerations have been taken in incorporating the KPIs for a team's incentive calculation. First consideration is each KPI's current contribution towards the respective department's performance. Second is each KPI's level of performance of current period compared to target.

$$KPI_{is} = \sum_{m=1}^{M} \sum_{c=1}^{C} \beta_{cm} KPI_{cm,Act.} \quad (4)$$

By using weighted KPI values to stochastic gradient descent algorithm, above relationship's β_{cm} can be solved for the current period. The ratios of $\frac{\beta_{11}}{\sum_{m=1}^{2} \beta_{cm}}$ and $\frac{\beta_{12}}{\sum_{m=1}^{2} \beta_{cm}}$ represents the selected employee team's (team 1) contribution of two KPIs towards department 1's performance for current period.

Ratio between actual and budgeted values of the two KPIs for the current period has been used to calculate the team's level of performance. Combining these two considerations on the team level performance, current period incentive (I_{11}) for an employee of team 1 is calculated as below.

$$I_{11} = M' \left[\left(\frac{\beta_{11}}{\sum_{m=1}^{2} \beta_{cm}} \frac{KPI_{11.Act}}{KPI_{11,Bud.}} \right) + \left(\frac{\beta_{12}}{\sum_{m=1}^{2} \beta_{cm}} \frac{KPI_{12.Act}}{KPI_{12,Bud.}} \right) \right] \quad (14)$$

4.6 Alternative Options for Target Setting

Corporate level targets and operational employee team level targets are used for the proposed incentive scheme. In general, periodic corporate level targets are planned in finance departments of the organizations. However, for operational team level, targets may not be prepared to each period separately. Therefore, it is important to have alternate options to calculate targets of the incentive scheme.
As the guideline for the targets, only the current standards and attainable standards have been considered. Out of the two, attainable standards are more preferable since it intends to have a reasonable improvement level of the team performance. (Kaplan, 2009; Chary, 1997)

$$KPI_{Budget} = KPI_{Average}(1 + improvement\ percentage) \quad (15)$$

$KPI_{Average}$ is determined with the use of few recent historic periodic values for KPIs. On top of the average value, target can be set to attain a reasonable improvement level. If the operational level indirect employees attain the target, they would entitle to get maximum incentive allocated for the respective KPI or in a case of partial achievement, respective partial achievement would be incentivized.

$KPI_{Budget} = KPI_{Average}$ or $KPI_{Budget} = KPI_{last\ period}$

For certain KPIs, operational level employee teams would have already reached the optimum acceptable level of performance. In such situations, current standards can be used to incentive calculation as given above.

4.7 Weighting the KPI Values to use in Gradient Descent

Stochastic gradient descent algorithm has been used in three instances of the incentive calculation in order to find the weights for performance contributions. Given below are the linear regression solved through the gradient descent.

$$KPI_{F,Act.} = \sum_{a=1}^{A} \theta_a KPI_a \quad (2)$$

$$KPI_a = \sum_{i=1}^{n} \sum_{s=1}^{S} \emptyset_{is} KPI_{is} \quad (3)$$

$$KPI_{is} = \sum_{m=1}^{M} \sum_{c=1}^{C} \beta_{cm} KPI_{cm,Act.} \quad (4)$$

It is required to weight the KPI training examples before they are used in gradient descent. Reason for that is, different values of KPIs may have different scales and it would have a high impact on theθ_a,$\emptyset_{is}$ or β_{cm} values (e.g.: values for KPI_{1s} may be given in thousands whereas values for KPI_{2s} may be given in decimals). Therefore it is important to bring different scales to a single scale before it is being used. As a solution, it is advised to weigh each KPI value from its average value of historic data.

4.8 Consideration on KPI Co-Relation to Organizational Performance

Some key performance indicators may have either positive or negative co-relation with the overall organizational performance. As an example, standard hours achieved can be considered as a positively co-related KPI with the company performance because, greater the standard hour’s achievement greater will be the opportunity to attain higher company performance. In contrast, a KPI like defect rate is having a negative co-relation with overall company performance because, greater the defect rate, lesser will be the opportunity to attain a higher company performance.

Attention has been paid in categorizing the possible KPIs based on their co-relation because it can affect the incentive scheme in two ways. If negatively co-related KPI data have been used in the gradient descent solving, respective gradient values will result minus values which creates problems in proceeding. In the case of actual to budget KPI ratios used in incentive scheme, it would result misleading results if a KPI is having a negative co-relation with company performance. Therefore as a solution, it is advised to convert the values of negatively co-related

KPIs to positively co-related values (1 - negative co-related KPI %) before they are processed to incentive calculation.

5. Discussion

As aforementioned, intention of a financial based incentive scheme is of two ways. It should simultaneously reward employees while motivating them and render good results to the employer. Latter intention with employer’s perspective is completely ignored if incentives are allocated to operational level indirect employees without considering their contribution towards overall organizational performance.

$$I_{ic} = \underbrace{\frac{\left(\left[X\left(\frac{KPI_{F,Act.}}{KPI_{F,Bud.}}\right)\left(\frac{\phi_{is}}{\sum_{i=1}^{n}\phi_{is}}\right)\right] + bal.\right)\left(\frac{1}{\sum_{i=1}^{n} N_i}\right)}{t}}_{\text{Employer's Perspective}} \underbrace{\left[\sum_{c=1}^{C}\sum_{m=1}^{M}\left(\frac{\beta_{cm}}{\sum_{c=1}^{C}\sum_{m=1}^{M}\beta_{cm}}\right)\left(\frac{KPI_{cm,Act.}}{KPI_{cm,Bud.}}\right)\right]}_{\text{Employees' Perspective}} \quad (1)$$

Two intentions of the employer and employee are considered under the proposed incentive scheme as highlighted above. Unless operational level employee’s department level contribution towards organizational performance is not considered, employer cannot easily recognize whether it is worthy to practice an incentive scheme.

Simplicity of an incentive scheme is a factor determines the understand ability and acceptability to an incentive scheme. There can be a possibility that the part representing employer’s perspective may be difficult in explaining to operational level indirect employees. As a solution to the above issue, following suggestion was made. It can be seen that the employer’s perspective part of the incentive scheme is a periodic constant amount for operational level indirect employees belong to the same department. Employees can be convinced that based on their department success and broader teamwork, the I' amount for the period varies. But their main concentration should be to achieve the complete department level I' amount for the period to their teams by fully achieving targets set for both the KPIs. If targets for both the KPIs are being achieved, they received a guaranteed minimum of the respective department’s incentive ceiling (I') for the period. In addition, it can be convinced to the employees that though the I' amount does not vary from a greater amount in worst scenario, based on their collective work contribution to the department level performance, it can be extended favorably.

$$I = I'\left[\sum_{c=1}^{C}\sum_{m=1}^{M}\left(\frac{\beta_{cm}}{\sum_{c=1}^{C}\sum_{m=1}^{M}\beta_{cm}}\right)\left(\frac{KPI_{cm,Act.}}{KPI_{cm,Bud.}}\right)\right] \quad (16)$$

6. Conclusion

A comprehensive mathematical model is developed for incentivizing the operational level indirect employees. Team based financial incentive scheme is developed by considering both employer and employees’ perspectives. Stochastic gradient descent algorithm is used for the mathematical formulation. The developed incentive scheme satisfies the factors that are essential to be considered when developing an incentive scheme for indirect employees as mentioned in section 2.5. Guidelines for the selection of appropriate KPIs for department level and operational level indirect employees is not included in this paper. Therefore, it is advisable for the users to establish proper KPIs with the use of a suitable performance measurement system. Maximum benefits of the proposed incentive model can be gained through the selection of appropriate KPIs, incentive period, employee teams and KPI target setting. As a future research direction, it is expected to test the developed model using actual data sets from a selected organization.

7. References

1. Allen, D. (2011), Levels of control. Available at: http://www.cimaglobal.com/Thought-leadership/Newsletters/Insight-e-magazine/Insight-2011/Insight-October-2011/Strategic-financial-management-levels-of-control/ (Accessed 20 October 2014)
2. Banfield, P. and Kay, R. (2012), *Introduction to human resource management*. Oxford University Press.
3. Burgess, S. and Ratto, M. (2003), the role of incentives in the public sector: Issues and evidence. *Oxford review of economic policy*, 19(2), pp.285-300.
4. Chary, S.N. (1995), *Theory and Problems in Production and Operations Management*. Tata McGraw-Hill Education.
5. Condly, S.J., Clark, R.E. and Stolovitch, H.D. (2003), the Effects of Incentives on Workplace Performance: A Meta-analytic Review of Research Studies 1. *Performance Improvement Quarterly*, 16(3), pp.46-63.
6. Cox, S.A. (2014) *Managing Information in Organizations: A Practical Guide to Implementing an Information Management Strategy*. Palgrave Macmillan.
7. Gordon, A.A. and Kaswin, J.L. (2010), Effective Employee Incentive Plans: Features and Implementation Processes.
8. Goyal, D.P. (2006), *Management Information Systems: Managerial Perspectives*. Macmillan.
9. Hamilton, B.H., Nickerson, J.A. and Owan, H. (2003), Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of political Economy*, 111(3), pp.465-497.
10. Hoffman, J.R. and Rogelberg, S.G. (1998), a guide to team incentive systems. Team Performance Management: *An International Journal*, 4(1), pp.23-32.
11. Holtmann, M. and Grammling, M. (2006), Designing staff incentive schemes to balance social and financial goals.
12. Holtmann, M. (2002), Principles for Designing Staff Incentive Schemes. Available from: http://www.microfinancegateway.org/gm/document-1.9.26843/22624_ST_Incentive_Design.pdf
13. Kaplan. (2009), Standard Costing [Online]. Available at: http://kfknowledgebank.kaplan.co.uk/KFKB/Wiki%20Pages/Standard%20Costing.aspx (Accessed 10 December 2014)
14. Kohn, A. (1993), why incentive plans cannot work. Harvard business review, 71(5).
15. McDougall, A. and Radvanovsky, R. (2008), *Transportation systems security*. CRC Press.

16. Perry, J.L., Engbers, T.A. and Jun, S.Y. (2009), Back to the Future? Performance-Related Pay, Empirical Research, and the Perils of Persistence. *Public Administration Review*, 69(1), pp.39-51.

17. Ratnayake et al., (2014), Application of Stochastic Gradient Descent Algorithm in Evaluating the Performance Contribution of Employees, *IOSR Journal of Business and Management (IOSR-JBM)*, e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 16, Issue 6(3), pp. 77-80.

About Our Authors

Vijitha Ratnayake: Senior lecturer in Work Study, Production Management, Lean Manufacturing, Production Planning & Operational research in the University of Moratuwa, Sri Lanka. She is carrying out research and consultancy works in the manufacturing industry in those areas.

Gayathri Perera: Master of Philosophy candidate in the University of Moratuwa, Sri Lanka. Research areas interested in are industrial engineering, human resources management, and Cellular Manufacturing Systems.

Big data Based Decision Making for the Selection of Celebrity Advertising Model: A Korean Case

Volume 24, *Number* 2
June 2018, *pp.* 153-172

Mina Woo
Jinwha Kim
Sogang University
(drminawoo@gmail.com)
(jinwhakim@sogong.ac.kr)
Gyoo Gun Lim
Hanyang University
(gglim@hanyang.ac.kr)

This study presents a method of big data analysis which enables to choose appropriate candidates for advertising model. For five famous Korean sports celebrities and eight commercial products, this study collected 5,844,167 data from Twitter and blogs. Employing the big data analysis, we found their mutual keywords by using the proposed keyword frequency analysis. Using big data analysis is very effective because the image suitability between products and advertising models can be checked in advance, while in the past enterprises had to pay high cost to the advertising model with a risk on the image suitability between model and product.

Keywords: Big Data, Celebrity Advertisement, Match-Up Hypothesis Theory, Image Suitability, Advertisement Model

1. Introduction

People in the modern era live in an environment where they can easily have access to diverse and plentiful information thanks to the distribution of the Internet, Wi-Fi, and smart phones. In other words, accessing information wherever and whenever we want has been possible because of the Information and Communication Technologies (ICTs), which even lead communication between real space and cyber space in our daily life. In particular, with the widespread use of social media, the number of social media users amounted to 1.96 billion worldwide in 2015. A new domain — the Social Networking Service (SNS) — was created for social media sites such as Facebook, Twitter and other blogs, successfully forming new networks among people.

The fact that monthly real time users of Twitter in 2015 surpassed 302 million and Facebook users surpassed 1.4 billion proves the importance of the SNS. The development of the SNS not only enlarged the significance of the netizen's opinions, but also enabled the SNS to advance into promotions and other analytical spaces of corporate marketing. The development of the SNS made it possible for enterprises to use SNSs as a method of diverse marketing and analysis. This is due to the fact that SNS advertising became a significant marketing space along with the continuous growth of mobile devices.

Amid the ongoing success of SNSs, big data analysis helps enterprises to make effective and scientific decisions by analyzing the immense amount of information that appears on SNSs. As a Research and Innovation program from 2014 to 2020, the

European Union invested more than 800 Euros to bring about more technological advancements in the field (Digital TV, 2014). However, though its importance is being recognized, big data research in the field of advertisement image fitness is insufficient for sports marketing.

Meanwhile, thanks to the development of digital media, modern people can directly or indirectly come across many advertisements every day. Due to the abundance of advertisements, enterprises are taking advantage of famous celebrities to enhance the advertising effect. The appearance of famous celebrities in advertisements not only promotes products and increases brand awareness, but also enhances the product reliability of consumers. Based on these reasons, companies invest huge amounts of money to hire celebrities for their commercials, causing a constant increase in the participation of sports celebrities in commercials.

Among the celebrities, sports celebrities have recently become loved by marketers because of their awareness and attention from the public. Since celebrity marketing is known as the most effective way to improve corporate brand awareness, using sports celebrities as endorsers is not a surprise. Actually, there are many cases where companies cast famous endorsers who possess all the necessary conditions such as awareness, reliability, and familiarity in their advertisements, and it has a large influence on product sales. According to the economic magazine Forbes, world famous golfer Tiger Woods earned 98.8% of his annual income from his advertisement endorser fee, which is about $50 million. Golfer Phil Mickelson and tennis player Roger Federer also earned 94.5% and 86.6% of their annual income from advertising. These facts directly show how much the enterprises invest in sports star advertising.

However, if companies wish to enhance the effectiveness of an advertisement, they should use famous celebrities suitable for the product in the advertisement. The effect of the advertisement is maximized when the celebrity has a high fitness with the product (Lee and Park, 2014). So, what methods can a company use to measure image fitness? The most commonly used methods are questionnaire analysis, professional insight, or intuitive judgment. However, questionnaire analysis has two critical weaknesses as an analysis method. First, the data collected through questionnaires is limited by the researcher's questions (Coughlan et al., 2009). Second, social desirable bias can affect the results of the questionnaire. In other words, a respondent may have a tendency to answer differently from his or her true opinion (Demaio, 1984), so the result of the questionnaire may not perfectly reflect the person's real thoughts.

Thus, this study utilized big data analysis through SNSs to find the image fitness between the sports celebrity and product in an advertisement. Using five famous sports celebrities and eight advertised products, this study collected 5,844,167 data points on Twitter and blogs from May 1, 2014 to May 1, 2015 and analyzed their mutual keywords. Using big data analysis has a large significance because the image fitness between products and advertising endorsers can be checked in advance, unlike in the past when enterprises had to pay the high cost of advertising in the commercial sector while taking a risk on the image fitness.

The following is the outline of this paper. Section 2 examines the studies about sports celebrities and match-up hypothesis theory. Section 3 explains the data analysis method, which uses big data. Section 4 discusses the results, and Section 5 states the conclusion of this study.

2. Literature Review

2.1 Sports Celebrity Advertisements

Today, a number of enterprises employ indirect communication using famous advertisement endorsers — a way to approach the public in a friendlier and amicable way — to increase their market value and promote their products. A famous celebrity's advertisements are more effective than any other marketing strategy because they can draw more consumer attention, improving consumer behavior and purchase intention. In other words, companies use celebrities who have public awareness as a reference to their products, thereby strengthening the advertisement effect (McCracken, 1989). This kind of public confidence is formed by the basic credibility of the celebrity with his or her professionalism and attractiveness (Hovland et al., 1953; McGuire, 1973; Sternthal et al., 1978), and many celebrity advertisements are based on this (Kamen et al., 1975; Friedman and Friedman, 1979; Mowen and Brown, 1981; Atkin and Block, 1983). Actually, it has been reported from advertisements for gas oil, color TVs, and disposable razors that awareness of the product and the brand rises when a company uses a celebrity instead of an ordinary person in an advertisement (Freiden, 1984; Kahle and Homer, 1985). Furthermore, studies have found that using a celebrity brought a positive response toward the product and improved the purchase intention (Mowen and Brown, 1981) and celebrity trustworthiness, expertise, and attractiveness effects on celebrity endorsement (Amos et al., 2015). This is especially true when the celebrity's original image and compatibility are suitable for the product. Because of this, the number of companies that take advantage of famous figures when advertising their newest products and improving their corporate image is gradually increasing.

Among celebrities, sports celebrities are most frequently used as advertisement endorsers. Sports celebrities bring attention to an advertisement and at the same time, transfer the celebrity's clean and pure image to the product or the brand, which can lead to a positive image. Establishing a positive image for the product or the brand through the advertisement is very important because it is related to the advertisement attitude, which is a consistently positive or negative response tendency toward the advertisement stimulus (Ajzen and Fishbein, 1980; Lutz et al., 1983; Chang et al., 2014). The advertisement attitude itself is a significant variable when measuring the advertisement effect, and using a sports celebrity as the advertisement endorser strengthens this trend. Considering these facts, we cannot ignore sports celebrities' contribution to intangible brand equity benefits (Bozman et al., 2015)

2.2 Match-up Hypothesis Theory

Using a celebrity as an advertisement endorser is theoretically based on the match-up hypothesis because the advertisement effect depends on the fitness between the endorser and the product. The match-up hypothesis was first presented in 1979 by Friedman and Friedman, who studied the fitness between advertisement endorsers and products. The match-up hypothesis means that suitable endorsers should be used for an advertisement to effectively deliver information to the consumers, thus strengthening the advertising effect (Kamins and Gupta, 1994). Therefore, the hypothesis assumes that the image of the advertisement endorser and the company or the product have to correspond with each other to enhance the consumer's memory and promote a positive response. In other words, positive effects from consumers can be obtained when the image of the advertisement endorser and the image of the corresponding brand are appropriate.

According to the match-up hypothesis, the experience of a measured fit — which occurs when the measured focus and provided stimulus match — enables positive value judgments by helping people feel confident about their current actions, thereby increasing their individual motivation to perform and immerse themselves in certain activities. In conclusion, a positive advertisement effect is possible only when the image of the advertisement endorser and the image of the product are appropriate.

Fitness is a concept from the brand extension, co-brand, or sponsorship to explain the awareness degree between the two different objects (Korchia et al., 2009). The cognitive theory defines this fitness concept in four different ways (Till and Busler, 2000). First, congruence means the category of the product and the celebrity advertised product's category have to be the same, but the endorser itself does not have to correspond with the brand (Kamins, 1990; Lynch and Schuler, 1994). On the other hand, fittingness is a concept about the suitability between the advertisement endorser and the product (Kanungo and Pang, 1973). While appropriateness asks the accordance of the image type (Solomon et al., 1992), consistency means the images between the advertisement endorser and the product have coherence (Walker et al., 1992). Table 1 shows four different concepts of fitness and their researchers.

The fitness between the advertisement endorser and the product is important because if the celebrity's image doesn't match up with the product, it not only damages the celebrity's credibility but also diminishes the advertising effect (Kamins et al., 1989; Kikati, 1987). Therefore, the match-up hypothesis has been supported by many researchers. Table 2 shows the research studies about the match-up hypothesis.

Table 1 *Four Different Concepts of Fitness*

Term	Researcher
Congruence	Kamins (1990), Lynch and Schuler (1994)
Fittingness	Kanungo and Pang (1973)
Appropriateness	Solomon, Ashmore, and Longo (1992)
Consistency	Walker, Langmeyer and Langmeyer (1992)

Friedman and Friedman (1979) researched the advertisement effects using three different types of advertisement endorsers – celebrity, expert and regular people. Based on the preliminary investigation, they chose vacuum cleaners as a functionally, physically, and financially high-risk product, jewels as a socially and psychologically high-risk product, and snacks as a low-risk product. By analyzing the subject's attitude towards the product and the advertisement and his or her purchase intention, they found that the fitness between the product and the advertisement endorser had a significant effect on the advertisement's credibility. A vacuum cleaner showed the best advertisement effects when it was advertised by an expert, as did the jewels with a celebrity and the snack with a general person.

Table 2 *Match-up Hypothesis Research Studies*

Researcher	Title	Summary
Cho and Baskin (2018)	It's a match when green meets healthy in sustainability labeling	The match-up between healthiness and sustainability drives consumer buying preferences and product perceptions.
Levi et al. (2017)	The Match-up Hypothesis Revisited: A social Psychological Perspective	The match-up relationship exists between advertising model's attractiveness and attractiveness-related product types.
Seiler and Kucza, (2017)	Source Credibility Model, Source Attractiveness Model and Match-Up-Hypothesis - An Integrated Model	Product fit has a positive effect on attitude towards the ad and in turn towards the brands and purchase intention.
Bahram et al. (2010)	Celebrity endorser influence on attitude toward advertisements and brands	If celebrity has a proper relationship with a product, celebrity image helps consumer to build the brand image
Koernig and Boyd (2009)	To catch a tiger or let him go: The match-up effect and athlete endorsers for sport and non-sport brands	Product-endorser fit can be related to a positive image and purchase intention toward the product
Till and Busler (2000)	The match-up hypothesis: Physical attractiveness, expertise, and the role of fit on brand attitude, purchase intent and brand beliefs	Fitness between a message of celebrity endorser and product makes an effective advertisement
Kamins (1990)	Celebrity and No celebrity Advertising in a Two-Sided Context	The fact that the celebrity's attractiveness is emphasized when he or she is paired with the product supports the match-up hypothesis
Kahle and Homer (1985)	Physical Attractiveness of the Celebrity Endorser	Physical attractiveness towards the product's brand is formed when a fascinating sports celebrity is used as an advertisement endorser
Friedman and Friedman (1979)	Endorser Effectiveness by Product Type	Celebrity creates different advertisement effects depending on the product
Kanungo and Pang (1973)	Effects of Human Models on Perceived Product Quality	The advertisement effects of the product depend on the advertisement endorser

It also has been proven that the fitness between the product and the advertisement endorser instills a favorable attitude towards the product. If the advertisement endorser fits the product, consumers experience a perceptual and attitudinal congruence, and therefore the psychological comfort not only leads to a favorable attitude (Kanungo and Pang, 1973), but also improves the purchase intent of the product (Kahle and Homer, 1985; Koernig and Boyd, 2009). In addition, a positive attitude toward celebrity endorser can influence on attitude toward brand through celebrity's image (Bahram, Zahra and Zahra, 2010). There are many examples that support the importance of fitness between the product and the advertisement endorser. In a car advertisement, a fascinating celebrity, Tom Selleck, was more effective as an advertisement endorser than a non-attractive celebrity (Kamins, 1990). Similar results were found in a study that examined the fitness between the celebrity's characteristics, such as attractiveness and professionalism, the product, such as pens, candy bars or energy bars, and its resulting advertisement effect (Till and Busler, 2000). In other words, to make an effective advertisement, the message that the celebrity gives and the product appeal have to be appropriate for each other. Latest studies have also found that there exist significant relationships between consumers and products under different contexts (Cho and Baskin, 2018; Levi at al., 2017; Seiler and Kucza, 2017), which signify the importance of match-up hypotheses theory.

2.3 Match-Up Hypothesis in a Sports Celebrity Advertisement

Advertisements that use sports celebrities are effective because sports celebrities not only possess public charm, but also communicate objectivity and credibility to consumers based on their extensive knowledge and experience in the field. This is why enterprises pay huge amounts for sports celebrities and anticipate advertisement effects. For example, Nike paid 1.44 billion dollars for Michael Jordan and Tiger Woods, and Gillette spent between 300 and 500 billion dollars on David Beckham (Edaily, 2015).

Companies continuously utilize sports celebrities as their advertisement endorsers because they have practical and instantaneous effects of the advertisements. In fact, according to AC Nielsen Korea's instant noodle market data, Ottogi, which selected Hyunjin Ryu as its advertisement endorser, came in second place by recording 18% and beating the previous runner-up, Samyang Foods (12.4%) in the first half of the instant noodle market's annual share (The Korea Economic Daily, 2015). Although the whole instant noodle market share decreased two % in 2014, Ottogi's Jin Ramen surpassed its original sales target by 10% and achieved 27.1% growth compared to the previous year (Chosunbiz 2015). Netizens who became aware of this report said the following things: "I want to have some instant noodle after seeing that," "Hyunjin Ryu is fascinating," and "The Hyunjin Ryu effect is enormous." These are some good examples that directly prove sports celebrity advertisement effects. These responses, as shown in the study based on questionnaire surveys, also prove that advertisement effects can be accurately and easily examined using big data.

Sports celebrity advertisements have the ability to equate a fresh and untainted image of sportsmanship, developing a corporate image that goes beyond the brand image. However, choosing an appropriate endorser is important because the opposite case can maximize negative images of the product or the brand no matter whom the company uses as a endorser.

However, most studies use questionnaires to demonstrate the effects of sports celebrity advertisement. Studies regarding advertisement fitness mostly make verifications through questionnaire surveys, not big data analysis.

Existing studies used questionnaire surveys to research advertisement fitness, but this study attempts to verify the correlation analysis between sports celebrities and products through big data analysis based on SNSs where the public's thoughts and feelings are shown as they truly are. Big data analysis is a tool for companies to perceive consumers' thoughts and understand their demands. Therefore, it can be actively used in the marketing field, making it possible to lead to the instantaneous effects of marketing when analyzing postings on Facebook and Twitter. In other words, corporate marketers are now able to directly analyze the public's opinions and thoughts through SNSs, where common words are frequently mentioned.

Big data analysis generally utilizes keyword frequency analysis based on SNSs. For example, big data has produced meaningful results such as predicting movies' box office success and incomes by performing a keyword frequency analysis on Twitter.

A study method that utilizes keyword analysis on SNSs takes place in a variety of fields. Big data studies based on frequency analysis also take place in all industry fields, including the education sector, the public service sector, and daily life where SNS data exist, and these results are assessed to be highly credible. Therefore, this study examines the fitness between the sports celebrities and the products using big data analysis based on the fact that advertisement endorsers have to be suitable to maximize the advertisement effect.

3. Methodology

3.1 Target Analysis of Sports Celebrities and Products

In order to select six sports celebrities, this study collected the following data: the revenue of advertisements and sports celebrities in 2014, the number of broadcasted data on Naver, the preference ranking of sports celebrity advertisement endorsers that was conducted by Gallup Korea's questionnaire survey, and Macromill Korea's survey, which targeted ordinary citizens and advertisers. As a result, we found the following results: 1st – Yuna Kim (estimated price of approximately 16.3 billion won with 64,215 votes), 2nd – Hyunjin Ryu (estimated price of approximately 10 billion won with 62,035 votes), 3rd – Yeonjae Sohn (estimated price of approximately 3 billion won with 17,889 votes), 4th – Taehwan Park (estimated price of approximately 5 billion won with 19,049 votes), 5th –Seonghoon Choo (estimated price of approximately 3 billion won with 19,45 votes), and 6th – Seongyong Ki (estimated price of approximately 6 billion won with 23,716 votes).

Based on these sports celebrities, data on consumers' feelings about products were collected from Twitter and blogs by conducting a keyword frequency analysis of eight different products: a refrigerator, massage machine, washing machine, air conditioner, telecommunications company, smartphone, food, and cosmetics from May 1, 2014 to May 1, 2015. Saltlux software framework, the program most widely used in this field, was utilized for data collection while controlling the current endorsement activities of above six sports celebrities.

As a result, by gathering 389,688 instances of keywords about Yuna Kim, 180,159 about Yeonjae Sohn, 1,149,94 about Hyunjin Ryu, 423,745 about Seongyong Ki, and 260,185 about Seonghoon Choo, a total of 1,368,771 keywords related to the sports celebrities were collected. In the product category, a total of 4,475,396 keywords – 295,198 keywords about the refrigerator, 1,977 keywords about the massage machine, 71,549 keywords about the washing machine, 427,300 keywords about the food, 310,662 keywords about the air conditioner, 1,025,583 keywords about the telecommunications company, 562,517 keywords about the smartphone, and 411,839 keywords about the cosmetics – were analyzed. Eventually, a total of 5,844,167 opinions from Twitter and blogs were collected and used in the data analysis. Since Taehwan Park was implicated in a doping incident, he was excluded and the analyses of the other five sports celebrities were used. Table 3 shows the number of analyzed keywords for each sports celebrity with a simple explanation, and Table 4 shows the number of analyzed keywords for each product.

Table 3 *The Number of Keywords for Each Sports Celebrity with an Explanation*

Sports celebrities		Explanation	Keywords
Yuna Kim		South Korean figure skater. 2009 and 2013 World champion. 2010 Olympic champion. 2014 silver medalist in ladies' singles.	389,688
Yeonjae Sohn		South Korean individual rhythmic gymnast. 2014 Asian Games all-around champion. 2014 World Championships 4^{th} in the all-around.	180,159
Hyunjin Ryu		South Korean professional baseball starting pitcher for the Los Angeles Dodgers of Major League Baseball.	1,149,942
Seongyong Ki		South Korean professional footballer. Central midfielder for the Premier League club Swansea City. Captain of the South Korean national team.	423,745
Seonghun Choo		Japanese mixed martial artist and judoka. 2001 Asian Championships gold medalist for South Korea. 2002 Asian Games gold medalist for Japan. Former K-1 HERO's Light Heavyweight Grand Prix Tournament Champion.	260,185
Total			1,368,771

Note: Images Retrieved from Google.com

Table 4 *The Number of Keywords for Each Product*

Product	**Keywords**
Refrigerator	295,198
Massage machine	1,977
Washing machine	71,549
Food	427,300
Air conditioner	310,662
Telecommunications company	1,025,583
Smartphone	562,517
Cosmetics	411,839
Total	4,475,396

3.2 Big Data Analysis Procedure

The big data analysis of this study was done in three stages after preprocessing the raw data.

In the first *extracting keywords* stage, the 100 most frequently mentioned keywords on Twitter and blogs about the five sports celebrities and the eight products were extracted in the order of frequency. Then, the frequency of those 2,600 keywords on Twitter and blogs was estimated. After arranging the extracted respective 200 keywords from Twitter and blogs about each sports celebrity and product, we analyzed the final 100 keywords.

In the second *mutual keywords* stage, keywords about each sports celebrity and product were visualized and commonly mentioned keywords were classified. These keywords were defined as 'mutual keywords,' which formulated a common image between the sports celebrities and products. Based on the collected data, this study made a new formula to calculate and recommend the most proper advertisement endorser of each product by weighting the frequency of mutual keywords. Table 5 and Table 6 show the formula.

Table 5 *Definition of the Image Fitness Formula Based on Weigh Frequency of Mutual Keywords I*

Definition	Calculation Formula
n	A total number of the rank (n= 100)
S_i	i^{th} mutual keyword of the sports celebrity(S) (i=1,…, n)
wS_i	Weight of the i^{th} mutual keyword of the sports celebrity(S) (i=1,…, n) $= \frac{\text{frquency of the mutual keyword}}{\text{sports celebrity(S)'s total number of mutual keyword}}$
P_j	j^{th} mutual keyword of the product(P) (j=1, …, n)
wP_j	Weight of the j^{th} mutual keyword of the product(P) (j=1, …, n) $= \frac{\text{frquency of the mutual keyword}}{\text{product(P)'s total number of mutual keyword}}$
m	The number of $S_i == P_J$
F_{ij}	$= wS_i + wP_j$ (where $S_i == P_j$)

In the last *calculating image fitness* stage, the percentage of common appearance of the keywords for both products and sports celebrities were measured in consideration of the overall frequency. Since the final calculated percentage was the frequency of commonly appearing keywords between sports celebrities and products, the percentage was defined as 'image fitness.'

Table 7 is a calculated example of image fitness. The total sum of 100 keywords about Yuna Kim was 389,688, and the sum of Smartphone keywords was 562,517.

Table 6 *Definition of the Image Fitness Formula Based on Weigh Frequency of Mutual Keywords II*

Definition	Calculation Formula
Maximage(S, P)	Max(F_{ij}) = The most suitable mutual keyword between the sports celebrity(S) and the product(P)
TFR(S, P)	The sports celebrity(S)'s total fit ratio of the product(P) = $\sum_{i}^{m} F_{ij}$ (i, j = 1, …, n)
Recommend(S, P)	Recommendation product(P) of the sports celebrity(S)

As a result, seven mutual keywords – advertisement, photo, birthday, sister, model (including advertisement model), and friend – were found. The total number of common key words was 96,030 (22,187 + 73,843), which proved that the image fitness between Yuna Kim and the Smartphone was 10.09%. Figure 1 depicts the keywords of Yuna Kim and the Smartphone with the seven mutual keywords shown at the center. The appearance of different keywords according to the type of sports celebrity proves that each celebrity has a different image, and that verifies the advantage of big data analysis.

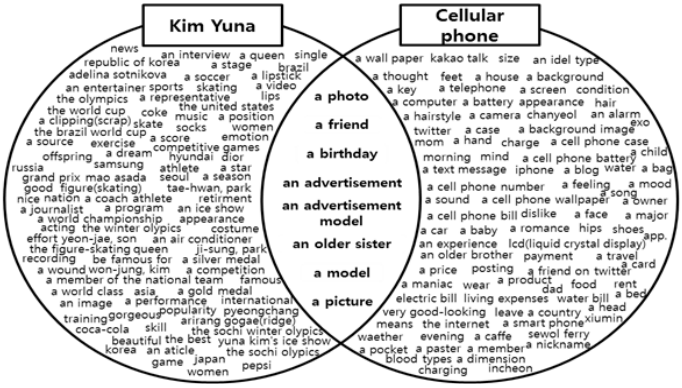


Figure 1 *Keyword Visualization of Yuna Kim and the Smartphone*

Table 8 shows the image fitness results for the five sports celebrities and eight products. In detail, Yeonjae Sohn showed high image fitness with the refrigerator (41.41%), massage machine (57.38%), and food (38.11%). Hyunjin Ryu showed high image fitness with the air conditioner (38.68%) due to his image as a powerful pitcher. Since Seonghoon Choo was frequently featured in the entertainment program 'Superman is Back' with his daughter Sarang Choo and displayed an image of a warm-hearted dad and family-oriented man, he had high image fitness with the laundry machine (41.56%), Telecommunications Company (28.30%), Smartphone (34.29%), and cosmetics (37.90%). The telecommunications company LG U+ even created an advertisement with the 'obsessed with daughter' character Seonghoon Choo.

Table 7 *Examples of the Frequency of Mutual Keywords and the Image Fitness between Yuna Kim and the Smartphone*

Rank	Sports Celebrity: Yuna Kim		Product: Smartphone	
	Keyword	**Frequency**	**Keyword**	**Frequency**
1*	Sports celebrity	29,284	Photo(28)	23,670
2	Figure skating	20,445	Background	19,439
3	Olympic	14,044	Key	18,341
10*	Advertisement(70)	5,967		
16*			Friend(96)	7,722
28*	Photo(1)	3,888		
42*	Birthday(51)	2,991		
51*			Birthday(42)	4,331
58*	Older sister(98)	2,610		
59*	Model(77, 78)	2,606		
70*			Advertisement(10)	3,613
77*			Advertisement model(59)	3,351
78*			Model(59)	3,344
79*	Picture(86)	2,249		
86*			Picture(79)	3,045
96*	Friend(16)	1,786		
98*			Older sister(58)	2,580
99	Kim Mu-seong	1.786	Charge	2,508
100	Correct answer	1.761	Inchon	2,505
Mutual keyword(*) sum		22,187(A)		73,843(B)
Total sum		389,688(C)		562,517(D)
Image fitness				10.09% (A+B)/(C+D)*100

Note: The Numbers in Parentheses Represent the Opposite Frequency Ranking about the Same Keyword

On the other hand, Yuna Kim generally had high image fitness, but no single product category stands out compared to the other sports celebrities. She showed high fitness with the smartphone (24.01%), but did not rank in first place despite her strong impression with the ‘Yuna Kim’ phone. In the case of Seongyong Ki, he turned out not to be suitable as an advertisement endorser since he lacked overall image fitness. The analysis results are discussed in further detail in the next section.

4. Results and Discussion

This study analyzed the fitness of sports celebrities with their respective products. For example, Figure 2 represents sports celebrities’ fitness with the refrigerator by percentage. As a result, we found that Yeonjae Sohn (41.4%) had the highest image fitness among the five celebrities in the refrigerator category, followed by Seonghoon Choo (40.5%), Hyunjin Ryu (11.4%), and lastly, Yuna Kim (6.7%). Yeonjae Sohn also had big differences from other sports celebrities, showing 57.38%

image fitness with the massage machine, 38.11 % with the food, and the second highest fitness with the cosmetics. Around the time when she earned 41.41% image fitness with the refrigerator, 57.38% with the massage machine, and 38.11% with the food, she gained a lot of popularity from the public.

Table 8 *Sports Celebrities' Image Fitness with Respective Brand and its Percentage (%)*

Percentage (%)	**Yeonjae Sohn**	**Yuna Kim**	**Seonyoung Ki**	**Hyunjin Ryu**	**Seonghoon Choo**
Refrigerator	5.18 (41.41%)	0.83 (6.67%)	0 (0.00%)	1.43 (11.47%)	5.06 (40.45%)
Massage machine	10.42 (57.38%)	1.72 (9.47%)	0.37 (2.06%)	0.63 (3.46%)	5.02 (27.63%)
Washing machine	6.87 (36.54%)	3.32 (16.71%)	0 (0.00%)	0.97 (5.19%)	7.81 (41.56%)
Food	4.07 (38.11%)	3.25 (30.46%)	0.71 (6.69%)	0 (0.00%)	2.64 (24.73%)
Air conditioner	2.18 (17.63%)	1.82 (14.75%)	0.61 (4.93%)	4.79 (38.68%)	2.97 (24.01%)
Telecommunications company	5.51 (20.77%)	6.04 (22.76%)	2.27 (8.56%)	5.20 (19.61%)	7.51 (28.30%)
Smartphone	7.80 (18.58%)	10.09 (24.01%)	0.43 (1.02%)	9.28 (22.10%)	14.40 (34.29%)
Cosmetic	7.64 (31.27%)	3.44 (14.06%)	1.68 (6.88%)	2.42 (9.90%)	9.26 (37.90%)

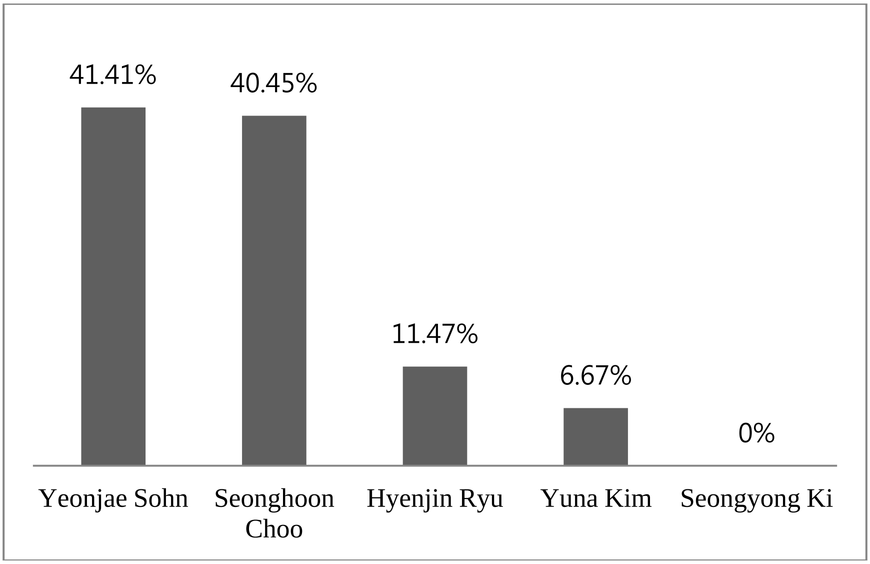


Figure 2 *Sports Celebrities' Image Fitness with the Refrigerator by Percentage (%)*

It is also necessary to make comparisons between the image fitness in the results of this study's big data analysis and the current situations of ongoing sports star advertisements in Korea. Therefore, we first compared Yeonjae Sohn and Yuna Kim, who have appeared in similar advertisements. Yeonjae Sohn had a high popularity level since she has a cute, pretty, and fresh image compared to Yuna Kim's mature image. Maeil Milk changed its advertisement endorser from Yuna Kim to Yeonjae Sohn and its sales increased by four times compared to the previous year. This result can be understood since Yeonjae Sohn is the green dietary life's honorary ambassador of Korea Agro-Fisheries & Food Trade Corporation, which aligns with the image she portrays as a cook in the advertisement. Furthermore, an advertising agency analyzed Yeonjae Sohn using the celebrity index, and the 2012 London Olympics Celebrity Endorsers Research Report Conference selected her as the No. 1 celebrity suited best for cosmetics commercials. In conclusion, the big data analysis explains that Yeonjae Sohn's confident performance and outward appearance have allowed her to surpass Yuna Kim in regards to public image; therefore, she was the sports celebrity with the highest preference.

Yeonjae Sohn had higher image fitness in cosmetics (31.27%) than that of Yuna Kim (14.06%). This is because although Yeonjae Sohn has appeared more frequently in advertisements with a similar image to Yuna Kim, she is less expensive for advertisements than Kim (Jang, 2014).

In the smartphone category, even though Yeonjae Sohn (18.58%) appeared in an LG OPTIMUS VN2 advertisement, Yuna Kim (24.01%) showed higher image fitness. Samsung's Anycall mobile phone used Yuna Kim as an advertisement endorser for quite a long time, and thus she is famous for the 'Yuna Kim' phone. The fact that Samsung's brand is stronger than LG in terms of smartphone awareness could also be a factor.

As for the category of air conditioner, although Yeonjae Sohn (17.63 %) and Yuna Kim (14.75 %) were the advertisement endorsers for the two most widely known air conditioners in Korea, Hyunjin Ryu proved to be the sports celebrity who was the most suitable to advertise an air conditioner (38.68 %). After examining corresponding keywords to figure out the reason for this result, the most important keyword turned out to be 'power.' In other words, Hyunjin Ryu had the highest image fitness with the refrigerator because of the public's assumption that power is an important attribute of an air conditioner, which corresponds with his image. Likewise, even though Yuna Kim and Yeonjae Sohn are the current advertisement endorsers, the big data analysis results could be helpful in selecting future advertisement endorsers since big data can disclose the public's unknown perceptions.

The reason why Seonghoon Choo showed a high fitness in the smartphone category is because many people talk about his smartphone on Twitter or blogs due to his frequent appearances on many entertainment shows. In fact, there was a question on the Naver search engine that said "Many people may want Seonghoon Choo to be an advertisement endorser. Which product commercial should he take?" and the answer was "mobile phone," indicating another reason for his high image fitness with the smartphone.

As the basis of this study finding, Samsung Mobile expected that the lovely image of Seonghoon Choo's sentimental appearances as a dad with his daughter Sarang Choo would generate and naturally deliver the wanted message to their brand targets. Thus, Samsung Mobile chose them to appear in a large scale Samsung America commercial that was advertised around the world along with famous world celebrities such as Messi (Kim, 2015).

Seonghoon Choo's highest image fitness with cosmetics may seem odd considering his previous job as a judo and mixed martial art player, but on the contrary, the public may see him as valuable. He has been the advertisement endorser for Coogi cosmetics since 2008 because he has a cold-hearted and masculine, yet smooth and warm-hearted image at the same time. His two different images are present: standing on the ring with his top off while at work and as his kind, normal self when not wearing his boxing gloves. Moreover, his appearance in the cosmetics product of Johnson & Johnson Korea along with his daughter was assessed to be another contributing factor that increased his image fitness with cosmetics (Tenasia, 2014). Also, there was only a narrow margin of 1.2% between him and Yeonjae Sohn, who was ranked number one for fitness as an advertisement endorser for Winia Mando's kimchi refrigerator 'Dimchae.' The increase of the public's interest toward Seonghoon Choo reflects the entertainment show's popularity and his position as an advertisement endorser (The Asia Economic Daily, 2015). In fact, he was selected as the Endorser of the Year at the 2008 Korea Advertisement Awards because he appeared in 10 commercials in diverse fields such as beer, the kimchi refrigerator, beverages, automobile, food, and cosmetics within a seven-month period (Maeil Business Newspaper, 2008). Therefore, this study's big data analysis result is meaningful in the sense that advertisers can observe this phenomenon closely and then select their future advertisement endorsers.

Even though Hyunjin Ryu was an instant noodle advertisement endorser, his image fitness did not stand out in the food category since instant noodles take up a very small portion of the category. However, if there was a single category of 'instant noodles,' the result would be different. Also, there was a time gap between his advertisement and this study's data collection. His instant noodle advertisement was broadcast in 2013, but this study collected its data between 2014 and 2015.

However, even though he did not stand out in this study's big data analysis, Hyunjin Ryu's instant noodle commercial demonstrates that sports celebrity commercials do have a significant effect. When Hyunjin Ryu was selected as the endorser for Jin Ramen in November 2013, its sales went up 30% and the instant noodle market share of Ottogi also went up (The Korea Economic Daily, 2015). This case shows how much a sports celebrity can influence a company's sales and the entire industry.

A comparative description of the results of this study's big data analysis was made with the current situation of the commercial market. As a result, certain sports celebrities who appeared in certain commercials significantly corresponded with this study's results, whereas other cases did not. Since disagreements signify the celebrities' possession of diverse images, companies need to look deeply into the big data analysis results.

5. Practical Implications

As sports have become an industry, sports celebrities with high awareness get a lot of attention from the public, which bolsters celebrity marketing as the most effective way to improve corporate brand awareness. In other words, since sports celebrities draw great attention from the general public, advertisements using them as endorsers have come to influence product sales greatly. Actually, companies are employing marketing strategies for the improvement of their corporate image by casting famous endorsers who possess all the necessary conditions such as awareness, reliability, and familiarity.

In this kind of environment, this study has huge significance because the image fitness between products and advertisement endorsers can be checked in advance through big data, unlike the past when enterprises had to pay endorsers in the commercial sector large fees and take a risk on the endorser fitness. Therefore, this study can definitely make a big contribution to the planning stage of advertisements by helping to figure out image fitness between the products and the sports celebrities. Through SNS keyword big data analysis of Twitter and blogs, now companies can understand the true opinions of the consumer.

When it comes to the study of image fitness, the questionnaire survey is the most basic and traditional method. However, the data from a survey is limited to answers to the researcher's questions. On the contrary, big data analysis has the advantage of understanding the public's ideas freely and unrestrictedly through SNSs.

Selecting and sponsoring proper sports celebrities who fit the product or company is a very important part of recent sports marketing. Therefore, this paper emphasizes the importance of decision making based on practical data that reflects the daily activities of the public over traditional questionnaire survey-based decision making. This can lead to data-based decision making and various forms of business endorsers. Also, this paper shows that a sports celebrity's reputation and image fitness with the product can be different from the result of the big data analysis. This means that if companies use big data analysis in their sports marketing, they can practically select and sponsor proper sports celebrities and save money while maximizing the advertisement effects using a scientific method.

As for the effectiveness of advertisements, this study can give practical assistance. Firstly, using big data makes it possible to hear the honest voice of the consumers' diverse opinions and demands, as opposed to the questionnaire survey method because SNSs are where people post their true opinions.

Secondly, this study shows that companies can prevent ineffective advertising expenses and obtain maximized commercial effects before they try to renew or create new contracts with sports celebrities by using big data analysis. If the image fitness between the advertisement endorser and the product can be evaluated before choosing an endorser, it can positively affect marketing practices. In particular, the result that shows consumers' loyalty and preference toward the product decreasing when image fitness does not match is noteworthy. That is because this result proves that the fitness between the image of sports celebrities and the image of products, brands, and companies are much more important than merely casting a famous sports celebrity as the endorser.

Lastly, SNS big data analysis will be very helpful when it comes to strategic decision making for corporate brand management and will offer competitive measures for establishing corporate brand marketing strategies. When big data analysis is applied to comprehend consumers' responses on SNSs, unnecessary advertising expenses will be reduced and strategies that create memorable advertising messages will increase.

6. Limitations and Future Studies

This study has a couple of limitations. First, the result of data mining analysis needs to be validated by alternative research methods. Considering the contribution and implication of this study, it is evaluated that the data mining method in this study is worthwhile and useful for the purpose of obtaining the immediate and reliable market data, but the validation of this method needs to be further investigated in the future. Second, this study has chosen only 5 sports celebrities and seven products within the Korean market. Due to the limited size of sample and the limited scope of industry within a specific nation, the proposed method needs to be conducted under diverse and different scenarios. However, even though the current study has aforementioned limitations, future studies will improve weaknesses and strength strengths in the light of evolving trend of big data analytics.

Future studies will find more profound results if they conduct long-term research projects with a greater number of keywords on products and sports celebrities that are not mentioned in this study. Furthermore, since the big data analysis can be supplemented by questionnaire surveys, such as those on detail purchase intention, that includes data that cannot be known when only using big data, a comparative study that estimates the similarities and the differences between big data and the actual survey is also necessary.

Meanwhile, this study demonstrates that big data analysis does not always correspond with the current situation and that the results may differ according to different search words. That is, big data analysis can discover a phenomenon that shows the mismatch of the current advertisement endorser and the product. Thus, if future studies use more specific and refined data, they will be able to find more prominent sports celebrity endorsers who draw active and positive responses from consumers.

One of the reasons behind the mismatched image between the result of the big data analysis and the ongoing advertisement market is the image of the sports celebrities themselves. Even though a sports celebrity forms a consistent image through games, that image can be changed when the sports celebrity appears in commercials or TV programs. For example, if sports celebrities display different or unknown aspects on TV programs, their image changes rapidly so different results are shown. From a theoretical perspective based on the concept of combination, a new sports celebrity image is created from the combination of different images that are generated by sports games and TV. In this case, the big data analysis shows new results regarding the new image because it reflects the rapidly changing SNSs. Therefore, more specific data refinement and analysis are necessary when carrying out big data analysis in this situation.

In conclusion, further studies are expected to produce more in-depth results by performing long-term searches based on a greater number of products and sports celebrity keywords that are not mentioned in this study. Additional investigations are necessary to verify the fitness of variable sports celebrities, a product's brand, and a product's category. Continuous research investing the differences depending on the type of consumer is also needed because it is an important variable to examine the substantive advertisement effect. In addition, if further research includes other variables such as advertisement exposure number, it could be possible to analyze the advertisement effect more objectively. This study also can be combined with a network analysis (Park and Lim, 2015). Lastly, a comparative study with an actual questionnaire survey needs to be done to supplement the results of the big data analysis. An actual questionnaire survey based on the results of this study could draw additional analyses that cannot be discovered solely through big data analysis, such as the detail purchase intention of the consumer.

Acknowledgements
This article was written based on Mina Woo's doctoral thesis.

7. References

1. Ajzen, I., and Fishbein, M. (1980), *Understanding attitudes and predicting social behaviour*, Prentice-Hall, Englewood Cliffs, NJ.
2. Amos, C., Holmes, G., and Strutton, D. (2015), "Exploring the relationship between celebrity endorser effects and advertising effectiveness- A quantitative synthesis of effect size", *International Journal of Advertising*, Vol. 27(No. 2), pp. 209-234.
3. Atkin, C., and Block, M. (1983), "Effectiveness of celebrity endorsers", *Journal of advertising research*, Vol. 23(No. 1), pp. 57-61.
4. Bahram R., Zahra S., and Zahra M. (2010), "Celebrity endorser influence on attitude toward advertisements and brands", *European Journal of Social Sciences*, Vol. 13(No. 3), pp. 399-407.
5. Bozman, C. S., Friesner, D., McPherson, M. Q., and Chase, N. M. (2015), "Intangible and tangible value: brand equity benefits associated with collegiate athletics", *International Journal of Sports Marketing and Sponsorship*, Vol. 16(No. 4), pp. 22-45.
6. Chang, Y., Jae Ko, Y., Tasci, A., Arai, A., and Kim, T. (2014), "Strategic match of athlete endorsement in global markets: an associative learning perspective", *International Journal of Sports Marketing and Sponsorship*, Vol. 15(No. 4), pp. 40-58.
7. Chosunbiz (2015), *Ottogi, Overtook Samyang Food and Consolidate Second Place in Instant Noodle Market*, Available from: http://biz.chosun.com/site/data/html_dir/2015/05/13/2015051303732.html.
8. Cho, Y. N., and Baskin, E. (2018). "It's a match when green meets healthy in sustainability labeling", *Journal of Business Research*, Vol. 86, pp. 119-129.
9. Coughlan, M., Cronin, P., and Ryan, F. (2009), "Survey research: Process and limitations", *International Journal of Therapy & Rehabilitation*, Vol. 16(No. 1), pp. 9-15.
10. DeMaio, T. J. (1984), "Social Desirability and Survey", Turner, C., and Martin, E. (Ed.), *Surveying subjective phenomena 2*, Russell Sage Foundation, pp. 257
11. Digital TV (2014), "EC and data industry make €2.5 billion 'big data' pledge", Available from: https://www.digitaltveurope.com/2014/10/13/europe-and-data-industry-make-e2-5-billion-big-data-pledge/.
12. Edaily (2015), *From Jordan to Yuna Kim' Sport, A Desire Economics*, Available from: http://starin.edaily.co.kr/news/NewsRead.edy?SCD=EB33&newsid=01092246609267568&DCD=A20402.
13. Freiden, J. B. (1984), "Advertising spokesperson effects-An examination of endorser type and gender on 2 audiences", *Journal of advertising research*, Vol. 24(No. 5), pp. 33-41.
14. Friedman, H. H., and Friedman, L. (1979), "Endorser effectiveness by product type", *Journal of Advertising Research*, Vol. 19(No. 5), pp. 63-71.

15. Hovland, C. I., Janis, I. L., and Kelley, H. H. (1953), *Communication and persuasion; psychological studies of opinion change*, Yale University Press: New Haven.
16. Jang, J. W. (2014), "One Hundred Million Won for One-Shot of a Drama, Celebrity Commercial Film Fees 'Over' One Billion Won", *Hankyung Business*, 28 March.
17. Kahle, L. R., and Homer, P. M. (1985), "Physical attractiveness of the celebrity endorser: A social adaptation perspective", *Journal of consumer research*, Vol. 11(No. 4), pp. 954-961.
18. Kamen, J. M., Azhari, A. C., and Kragh, J. R. (1975), "What a spokesman does for a sponsor", *Journal of Advertising Research*, Vol. 15(No. 2), pp. 17-24.
19. Kamins, M. A. (1989), "Celebrity and non celebrity advertising in a two-sided context", *Journal of advertising research*, Vol. 29(No. 3), pp. 34-42.
20. Kamins, M. A. (1990), "An investigation into the "match-up" hypothesis in celebrity advertising: When beauty may be only skin deep", *Journal of advertising*, Vol. 19(No. 1), pp. 4-13.
21. Kamins, M. A., and Gupta, K. (1994), "Congruence between spokesperson and product type: A matchup hypothesis perspective", *Psychology & Marketing*, Vol. 11(No. 6), pp. 569-586.
22. Kamins, M. A., Brand, M. J., Hoeke, S. A., and Moe, J. C. (1989), "Two-sided versus one-sided celebrity endorsements: The impact on advertising effectiveness and credibility", *Journal of Advertising*, Vol. 18(No. 2), pp. 4-10.
23. Kanungo, R. N., and Pang, S. (1973), "Effects of human models on perceived product quality", *Journal of Applied Psychology*, Vol. 57(No. 2), pp. 172-178.
24. Kikati, J. C. (1987), "Celebrity advertising: a review and synthesis", *International Journal of Advertising*, Vol. 6(No. 2), pp. 93-105.
25. Kim, M. J. (2015), "Choo Father and Daughter, American Advertisement Cast", *Ilgansports*, 29 April.
26. Koernig, S. K., and Boyd, T. C. (2009), "To catch a tiger or let him go: The match-up effect and athlete endorsers for sport and non-sport brands", *Sport Marketing Quarterly*, Vol. 18(No. 1), pp. 25-37.
27. Korchia, M., Fleck, N., and Le Roy, I. (2009), *Celebrities in advertising: looking for congruence or for likability?*, Paris Dauphine University.
28. Lee, J. G., and Park, J. (2014), "The effects of endorsement strength and celebrity-product match on the evaluation of a sports-related product: the role of product involvement", *International Journal of Sports Marketing and Sponsorship*, Vol. 16(No. 1), pp. 50-69.
29. Levi, E., Varnali, K., and Tosun, N.B. (2017), "The Match-Up Hypothesis Revisited: A Social Psychological Perspective", *International Journal of Communication*, Vol. 11, pp. 278-230.
30. Lutz, R. J., MacKenzie, S. B., and Belch, G. E. (1983), "Attitude toward the ad as a mediator of advertising effectiveness: Determinants and consequences", *Advances in consumer research*, Vol. 10(No. 1), pp. 532-539.
31. Lynch, J., and Schuler, D. (1994), "The matchup effect of spokesperson and product congruency: A schema theory interpretation", *Psychology & Marketing*, Vol. 11(No. 5), pp. 417-445.

32. Maeil Business Newspaper (2008), *KTF `Show` Win the Korea Advertising Grand Prize Twice*, Available from: http://news.mk.co.kr/newsRead.php?year=2008&no=640773.
33. McCracken, G. (1989), "Who is the celebrity endorser? Cultural foundations of the endorsement process", *Journal of consumer research*, Vol. 16(No. 3), pp. 310-321.
34. McGuire, W. J. (1973), "The yin and yang of progress in social psychology: Seven koan", *Journal of Personality and Social Psychology*, Vol. 26(No. 3), pp. 446-456.
35. Mowen, J. C., and Brown, S. W. (1981), "On explaining and predicting the effectiveness of celebrity endorsers", *NA-Advances in Consumer Research*, Vol. 8, pp. 437-441.
36. Park, B. E., and Lim G. G. (2015), "A Study on the Impact Factors of Contents Diffusion in Youtube using Integrated Content Network Analysis, *Journal of Intelligent Information Systems*, Vol. 21(No. 3), pp. 41~58.
37. Seiler, R., and Kucza, G. (2017), "Source Credibility Model, Source Attractiveness Model and Match-Up-Hypothesis–An Integrated Model", *Economy & Business Journal*, Vol. 11(No. 1), pp. 1-15.
38. Solomon, M. R., Ashmore, R. D., and Longo, L. C. (1992), "The beauty match-up hypothesis: Congruence between types of beauty and product images in advertising", *Journal of advertising*, Vol. 21(No. 4), pp. 23-34.
39. Sternthal, B., Dholakia, R., and Leavitt, C. (1978). "The persuasive effect of source credibility: Tests of cognitive response", *Journal of Consumer research*, Vol. 4(No. 4), pp. 252-260.
40. Tenasia (2014), *Choovely, Sarang Choo, Selected as a Cosmetic Advertisement Endorser*, Available from: http://tenasia.hankyung.com/archives/206913.
41. The Asia Economic Daily (2015), *Advertisement Effect 'Grew Up' Thanks to the Triplets*, Available from: http://view.asiae.co.kr/news/view.htm?idxno=2015060611311530616.
42. The Korea Economic Daily (2015), *Star Marketing 'Hyunjin Ryu Effect' Market Share 'Change-Up*, Available from from: http://www.hankyung.com/news/app/newsview.php?aid=2015012634721
43. Till, B. D., and Busler, M. (2000), "The match-up hypothesis: Physical attractiveness, expertise, and the role of fit on brand attitude, purchase intent and brand beliefs", *Journal of advertising*, Vol. 29(No. 3), pp. 1-13.
44. Walker, M., Langmeyer, L., and Langmeyer, D. (1992), "Celebrity endorsers: Do you get what you pay for?", *Journal of Consumer Marketing*, Vol. 9(No. 2), pp. 69-76.

About Our Authors

Mina Woo is a lecturer of Graduate School of Business at Sogang University in Seoul, Korea. She earned her marketing PhD in the Department of Global Service Management at Sogang University. Her major research interests are IT service management, big data analytics, and service marketing in Airline Service industry. She has publication in *Technological Forecasting and Social Change*, *Asia Pacific Journal of Tourism Research* (APJTR), *Sustainability*, *Indian Journal of Science and Technology* and *Journal of Korea Service Management society*

Jinwha Kim is a professor in the School of Business, Sogang University, S. Korea. His area of specilty is Management Information System. He was an assistant professor in the School of Business, Oklahoma State University. He got his Masters degrees in Business(MIS) and Computer Sciences from University of Wisconsin at Madison. He also got Ph.D in Business(MIS) from same university. His research area covers big data, data mining and customer relations management. He was a president of Korean Intelligent Systems, a vice president of E-Business Societies and a vice president of Big Data Association. He has more than 50 journal publications in bigdata and 11 books on bigdata and information technologies.

Gyoo Gun Lim, corresponding author, is a professor of MIS at the Hanyang University Business School. Before joining Hanyang University, he was an associate professor at Sejong University and a researcher with Samsung Electronics and Korea Telecom. He was a Fulbright visiting scholar at the Hass Business School at U.C. Berkeley in 2012. He received an award from the Korea Ministry of Information and Communication for his contribution to Korea SW industry in 2007 and an award from the Korea Ministry of Knowledge Economy for his contribution to Korea IT innovation in 2009. His current research interests include IT service, innovative business models, e-business, and intelligent information systems.

Online access of International Journal of Operations and Quantitative Management

Welcome to Online access of International Journal of Operations and Quantitative Management

- Please visit http://www.ijoqm.org
- Enter your Member/Subscription login data as,
  - Email: email address (as provide in subscription form)
  - Password: Already sent through email
- Please change your password using Edit Profile after logging in.
- You can make changes in your profile so that we have your current and complete information (such as address) with us.
- If there is any change in your email address or you have not supplied your email address, you will not be able to access the journal online. Please get your email address corrected or supply us your email address so that you can access IJOQM online.
- Click Past Articles and go through PDF version of all past articles.

Questions and Information

Please e-mail to managing-editor@ijoqm..org

Enjoy International Journal of Operations and Quantitative Management anywhere and everywhere!!!

Sixteenth
AIMS International Conference on Management
(AIMS-16)

Venue

Symbiosis Institute of Management Studies
Pune

Dates

January 3-5, 2019

Abstract Submission Date: October 31, 2018

Information: www.aims-international.org/aims16

Contact: aims16@aims-international.org