Transportation Research Part E 66 (2014) 1–22

Contents lists available at ScienceDirect

Transportation Research Part E

journal homepage: www.elsevier.com/locate/tre

A hybrid multi-objective approach to capacitated facility location with flexible store allocation for green logistics modeling

Irina Harris [a,*], Christine L. Mumford [b,1], Mohamed M. Naim [a,2]

[a] *Cardiff Business School, Cardiff University, Aberconway Building, Column Drive, Cardiff CF10 3EU, UK*
[b] *School of Computer Science & Informatics, Cardiff University, Queen's Buildings, 5 The Parade, Cardiff CF24 3AA, UK*

ARTICLE INFO

Article history:
Received 25 July 2013
Received in revised form 27 November 2013
Accepted 27 January 2014

Keywords:
Facility location–allocation problem
Green logistics
Multi-objective optimization
CO_2 emissions

ABSTRACT

We propose an efficient evolutionary multi-objective optimization approach to the capacitated facility location–allocation problem (CFLP) for solving large instances that considers *flexibility at the allocation level*, where financial costs and CO_2 emissions are considered simultaneously. Our approach utilizes suitably adapted Lagrangian Relaxation models for dealing with costs and CO_2 emissions at the allocation level, within a multi-objective evolutionary framework at the location level. Thus our method assesses the robustness of each location solution with respect to our two objectives for customer allocation. We extend our exploration of selected solutions by considering a range of trade-offs for customer allocation.

1. Introduction

One of the most important strategic issues for many businesses is where to site various storage and service facilities, and which facility should serve a particular customer. Usually there are restrictions that determine potential locations for facilities, and geographical, planning or financial constraints will influence a facility's capacity for storing goods and/or serving customers. Traditional capacitated facility location problem (CFLP) formulations aim to minimize the overall costs while maintaining a specific service level (Klose and Drexl, 2005; Nozick and Turnquist, 1998). Progressively more, businesses are under pressure to reduce the environmental impact from their operations (Beamon, 1999; Farahani et al., 2010; Lee and Dong, 2008; Sheu et al., 2005) where reporting emissions become a part of corporate social responsibility. The motivation for our work stems from our view that, with recent environmental concerns and high levels of commercial competition, companies could benefit from using multi-objective optimization (MOO) techniques when designing their distribution networks, to give extra flexibility to deal with key objectives simultaneously (Harris et al., 2011a); for example, minimizing the environmental impact of their business, while improving customer service levels at the same time as they are reducing cost or maximizing profits. Instead of a single solution (usually optimized on cost), multi-objective optimization techniques can offer a decision maker a choice of trade-off solutions, providing sufficient options to give him/her the power to make an informed choice that balances ALL the important objectives. It may be possible, for example, to greatly reduce energy

* Corresponding author. Tel.: +44 (0) 29 20874447.
E-mail addresses: HarrisI1@cardiff.ac.uk (I. Harris), C.L.Mumford@cs.cardiff.ac.uk (C.L. Mumford), NaimMM@cardiff.ac.uk (M.M. Naim).
[1] Tel.: +44 (0) 29 20875305.
[2] Tel.: +44 (0) 29 20874635.

http://dx.doi.org/10.1016/j.tre.2014.01.010
1366-5545

consumption or waste gas emissions whilst incurring a very small increase in economic costs. Such compromise solutions can be easily missed using traditional single objective methods.

The CFLP provides a popular model for various distribution networks, and involves making decisions at two levels: (1) which facilities to open from a set of potential sites, and (2) which customers to assign to which open facilities. Level 1 is usually considered as a *strategic decision*, because investing in storage facilities will tend to require large financial resources. On the other hand, once the facilities are in place, it is common practice to periodically reassess the allocation of customers at Level 2, which is a *tactical decision*. Although an initial assignment of customers to facilities at Level 2 is an essential part of making the strategic decision at Level 1, a single optimal solution obtained in this way is very likely to lack flexibility, and small changes made to the allocation of customers to facilities may produce unwanted consequences in terms of cost increases or quality of service reductions, for example.

The CFLP requires that all constraints are complied with, for example, all customer demand should be met and no capacity constraint should be violated (i.e., demand on a particular facility must not outstrip supply). Traditionally the design of a distribution network is driven by a need to reduce costs or maximize profit, and for this reason formulation is generally based on single objective optimization with all other potential objectives, such as customer service levels, modeled as extra constraints, so that a certain minimum level (e.g., of customer service) has to be maintained.

In the present paper, we are concerned with a *single source facility location*. We propose a novel approach for solving the CFLP, based on hybridizing a multi-objective evolutionary algorithm (MOEA) with a Lagrangian Relaxation (LR) approach that extends our recently published work (Harris et al., 2011b). The MOEA is used at Level 1 to determine which facilities to open, and the LR solves the generalized assignment problem (GAP) to allocate customers to serving facilities. The multi-objective CFLP could be solved using mixed integer programming software with a weighted sum approach to tackle levels (1) and (2) simultaneously but those approaches do not consider flexibility at the allocation level or the robustness of the solution over a period of time. It is our opinion that the location problem should be solved with flexibility and robustness in mind. We choose SEAMO2 (Simple Evolutionary Multi-objective Optimization 2) (Mumford, 2004; Valenzuela, 2002) as our MOEA to determine which facilities to open. For the allocation of customers to open facilities, we use a LR technique, to quickly assess the robustness of each location solution with respect to the GAP and our two objectives: cost and CO_2 emissions.

The rest of the paper is organized as follows. Section 2 reviews existing literature pertinent to our aim, with Section 3 describing the research method we employ. Section 4 introduces a problem formulation for our multi-objective CFLP. Sections 5 and 6 present our multi-objective optimization framework with a description of the elitist evolutionary multi-objective algorithm SEAMO2 and our experimental method. In Section 7, we present our Lagrangian Relaxation technique and Section 8 describes our test data. In Section 9 the tuning of SEAMO2 is described and this is followed by testing our LR technique on the published data in Section 10. Section 11 discusses our results and finally, we summarize our findings and provide suggestions for future research in Section 12.

2. Literature

A detailed overview of facility location formulations and solution techniques is presented in Daskin (1995), Drezner and Hamacher (2002), and Owen and Daskin (1998). Klose and Drexl (2005) review some contributions to the current state-of-the-art in facility location models for distribution system design, and provide a classification for facility location problems (FLPs) which considers different aspects of the design and includes: *discrete vs continuous* location models, *uncapacitated vs capacitated* models, *single-source vs multiple-source*, *static vs dynamic* models and others.

Lagrangian Relaxation (LR) techniques are leading methods for solving large CFLP problems (Barcelo and Casanovas, 1984; Beasley, 1988; Holmberg et al., 1999; Avella et al., 2009). Other techniques, such as approximation algorithms and metaheuristic approaches are also applied to solving those model formulations. Fisher (1981, 1985) presents separate traditional LR relaxation formulations (e.g. relaxing an assignment (allocation) constraint and a capacity constraint) to the GAP, where he discusses a number of applications of the technique in different areas. Jornsten and Nasberg (1986) propose a new LR approach based of a reformulation of GAP by introducing new substitution decision variables and new constraints. They show that the bounds from the Lagrangian dual of their approach are at least as strong as the bounds from traditional LR approaches. In the present paper, we present a mathematical formulation of a LR technique where we relax two capacity constraints simultaneously: the number of cases and the number of stores, in contrast to the traditional Lagrangian Relaxation where typically only the number of cases (or similar measure of demand quantity) is relaxed. We present in Section 7, a detailed description of the LR heuristic focusing on the techniques which are used to obtain upper and lower bounds.

Although the CFLP is NP-Hard (Mirchandani and Francis, 1990), Mixed Integer Programming (MIP) packages, such as CPLEX, can be used very effectively for small and even moderate sized instances to solve them to optimality. In addition, the software allows the user to adjust the gap tolerance to find good solutions to larger instances, but computational times depend on the complexity of the model formulation and actual values used. In this paper, we consider a multi-objective optimization problem for larger size instances, which requires a large number of repeated evaluations for each of the objective functions to obtain a set of trade-off solutions which demonstrates a need for efficient heuristic/metaheuristic approaches. In addition, we aim to explore a full range of solutions available to the decision-maker. This is in contrast to classical approaches solved using MIP packages, in which the multiple objectives are essentially reduced to a single objective, either by remodeling some objectives as constraints, or by combining the objectives into a single value using specific weights assigned to the objective functions. Thus classical methods do not allow such a free exploration of trade-off solutions.

2.1. Multi-objective optimization related to CFLP

Multi-objective optimization problems are common in the real world (Gen and Cheng, 1997; Deb, 2001) and such problems are characterized by optimum sets of alternative solutions, known as *Pareto sets*, rather than by a single optimum (Deb, 2001; Konak et al., 2006). Pareto-optimal solutions are non-dominated solutions in the sense that it is not possible to improve the value of any one of the objectives, in such a solution, without simultaneously degrading the quality of one or more of the other objectives in the vector. Evolutionary algorithms (EAs) are ideally suited to multi-objective optimization problems because they produce many solutions in parallel through imitating of a natural evolutionary progression (Gen and Cheng, 1997; Deb, 2001). EAs work with a different population of solutions in each iteration and produce a number of viable alternatives, rather than converging to a single optimal solution. Many successful applications of evolutionary algorithms within single objective setting have been published in the academic literature for a wide number of applications (Bäck, 1996; Michalewicz, 1996) and in this section of the paper we will present literature related to multi-objective setting with the focus on multi-objective evolutionary algorithms (Deb, 2001; Florios et al., 2010).

The literature related to multi-objective optimization within a facility location/allocation context can be classified into: (1) research that applies MOEAs and (2) other multi-criteria approaches, known as classical/preference based methods (Deb, 2001). In classical methods, a multi-objective problem may be converted into a single objective formulation using a weighted sum approach, or some objectives may be reformulated as constraints, for example. The Analytic Hierarchy Process (AHP) is well known technique and has been used to assign different weighting to quantitative and qualitative measures (Meade and Sarkis, 1998; Min and Melachrinoudis, 1999). To help put our work in context, Table 1 presents selected academic papers related to strategic network design (with CFLP) in terms of their use of multiple objectives (traditional and environmental) and their solution techniques (i.e., whether they use MOEAs). There is a shortage of published research on multi-objective optimization for the FLP using EAs or considering environmental objectives.

As mentioned above, research into location design with MOEAs application is still rather sparse and limited. The reader is referred to Deb (2001) for a detailed account of evolutionary multi-objective optimization where a number of evolutionary approaches are discussed. For example, the Elitist Non-Dominated Sorting Genetic Algorithm II (NSGA-II) has recently attracted a small number of publications for a traditional network problems (e.g. *min cost* and *max service level*) (Ding et al., 2006; Villegas et al., 2006; Bhattacharya and Bandyopadhyay, 2010; Liao et al., 2011a,b; Javanshir et al., 2012; Cheshmehgaz et al., 2013; Hiremath et al., 2013). For example, Villegas et al. (2006) present three algorithms (Non-dominated Sorting Genetic Algorithm, Pareto Archive Evolution Strategy and an algorithm based on mathematical programming) to minimize

Table 1
Academic studies related to MOO and network design.

References	Objectives				MOEA	Flexibility
	Traditional			Green		
	Cost[a]	Service	Other			
Ding et al. (2006)	✓	✓			✓	
Altiparmak et al. (2006)	✓	✓	✓			
Villegas et al. (2006)	✓	✓			✓	
Xu et al. (2008)	✓	✓			✓	
Bachlaus et al. (2008)	✓		✓		✓	
Bhattacharya and Bandyopadhyay (2010)	✓				✓	
Liao et al. (2011a)	✓	✓				
Cardona-Valdés et al. (2011)	✓	✓				
Paksoy et al. (2012)	✓					
Javanshir et al. (2012)	✓	✓			✓	
Hiremath et al. (2013)	✓	✓	✓		✓	
Cheshmehgaz et al. (2013)	✓	✓			✓	
Latha Shankar et al. (2013)	✓		✓		✓	
Khoo et al. (2001)	✓		✓	✓		
Hugo and Pistikopoulos (2005)			✓	✓		
Quariguasi Frota Neto et al. (2008)	✓			✓		
Pati et al. (2008)	✓			✓		
Bojarski et al. (2009)			✓	✓		
Wang et al. (2011)	✓			✓		
Harris et al. (2011b)	✓			✓	✓	
Pishvaee and Razmi (2012)	✓			✓		
Chibeles-Martins et al. (2012)	✓			✓		
You et al. (2012)	✓		✓	✓		
Current paper	✓			✓	✓	✓

[a] e.g. Cost of total supply chain; total cost of warehouse.

cost and maximize coverage for the uncapacitated facility location problem. Ding et al. (2006) adapted NSGA-II for multi-objective decision makers as part of their toolbox to support the design of supply chain networks. Liao et al. (2011a) propose a multi-objective model that integrates inventory decisions into a typical facility location model that minimizes the total cost and maximizes customer service level (volume fill rate and responsive level) with computational results presented for a problem with 50 buyers and 15 potential DCs. Javanshir et al. (2012) present a multi-objective particle swarm optimization and NSGA-II algorithm as a solution technique for a multi-stage supply chain network with two objectives: minimizing the total cost and maximizing total responsiveness. Xu et al. (2008) apply a spanning tree-based genetic algorithm using the Prufer number representation to design a network that minimizes total costs and maximizes customer service level under a condition of random fuzzy customer demand and transportation costs between facilities. Altiparmak et al. (2006) present a solution procedure based on genetic algorithms for a single product, multi-stage network design problem that minimizes the total cost, maximizes customer service and minimizes equity of the capacity utilization ratio.

Other approaches that are related to the behavior of swarms have also recently attracted academic attention. Bachlaus et al. (2008) propose a hybrid taguchi-particle swarm optimization algorithm that aims to minimize the cost (fixed and variable) and maximizes the plant and volume flexibility for a strategic design. Latha Shankar et al. (2013) present a multi-objective hybrid particle swarm optimization algorithm to solve a four-echelon network design to minimize the combined facility location and shipment costs subject to a requirement that maximizes customer demands. A hypothetical case study (3 suppliers, 5 plants, 6 DCs and 7 customer zones) is used to illustrate application of their approach.

Increasingly, the UK and other nations throughout the world are recognizing the importance of reducing the environmental impact of their operations (e.g., see Beamon, 1999; Farahani et al., 2010), and modern multi-objective optimization techniques (Deb, 2001) offer the flexibility to find good compromise solutions that balance the many important objectives, so that a decision maker has a choice of trade-off solutions without the need to make *a priori* decisions regarding the relative importance of the various objectives. From our literature review, we identified a small number of academic papers (e.g. Khoo et al., 2001; Hugo and Pistikopoulos, 2005; Quariguasi Frota Neto et al., 2008; Pati et al., 2008; Bojarski et al., 2009; Farahani et al., 2010; Wang et al., 2011; Harris et al., 2011b; Chibeles-Martins et al., 2012; Pishvaee and Razmi, 2012; You et al., 2012) focusing on traditional network design with multi-objective and environmental aspects, where a facility is serving a distribution center or warehouse. For example, Pati et al. (2008) use goal programming to balance economic and environmental goals for increasing wastepaper recovery in a paper recycling logistics system. Khoo et al., 2001 use a simulation approach in modeling a supply chain concerned with the distribution of raw aluminum metal. Quariguasi Frota Neto et al. (2008) describe the reorganization of a European pulp and paper logistic network where they assess the environmental impact using an environmental index (life cycle analysis) proposed in Bloemhof et al. (1996).

Bojarski et al. (2009) present the multi-objective MILP algorithm to solve a multi-period supply chain that optimizes two objective functions: NPV and environmental impact using a life cycle analysis, IMPACT2002+, methodology. Pishvaee and Razmi (2012) developed the multi-objective fuzzy mathematical programming model with an interactive fuzzy solution approach, based on the ε-constraint method, and the possibilistic programming approach that considers economic and environmental objectives (LCA method) in the design of reverse and forward supply chain networks. Chibeles-Martins et al. (2012) present a metaheuristic approach based on the simulated annealing methodology for a bi-objective mixed linear programming model with economic and environmental objectives. Wang et al. (2011) apply a normalized normal constraint method to solve a multi-objective mixed-integer formulation for the supply chain network problem that optimizes total costs and total CO_2 emissions of the supply chain.

The above literature review, together with Table 1, indicates that there is very limited research where application of evolutionary algorithms has been explored as part of the location analysis, especially that considers *flexibility* at the tactical customer allocation level; the *robustness* of the solutions over period of time where sustainability is a concern and environmental impact has to be considered as an objective function. The aim of the current paper is to consider environmental impact, in terms of CO_2 emissions as well as analyzing the robustness of generated solutions as part of the multi-objective framework. The purpose and potential novelty of the presented work is emphasized further in Section 3.

3. Method

The proposed framework for solving multi-objective CFLP applies an elitist evolutionary multi-objective algorithm, SEAMO2 (Simple Evolutionary Algorithm for Multi-objective Optimization) (Mumford, 2004; Valenzuela, 2002), to generate a set of trade-off solutions at the location stage. This stage determines which facilities will be open and which will be closed. We utilize a Lagrangian Relaxation technique (Fisher, 1981, 1985; Jornsten and Nasberg, 1986; Ghiani et al., 2004) to perform the allocation of customers to depots at tactical Level 2, which is described in detail in Section 7.

We build on the multi-objective work carried out on the uncapacitated facility location problem (UFLP) in Harris et al. (2009). The present paper also expands on our recently published work on the capacitated facility location problem (Harris et al., 2011b). Harris et al. (2011b) produced some encouraging results for the CFLP, balancing economic cost and CO_2 emissions. For simplicity, however, the allocation of customers to facilities was optimized using Lagrangian Relaxation on the basis of *economic cost only*, ignoring the fact that each and every location solution, lends itself to the multi-objective allocation of customers, as illustrated in Fig. 1. The present paper extends the Lagrangian Relaxation model that we use to assign customers to facilities, in three ways: (1) by relaxing two capacity constraints, instead of one, (2) by optimizing the allocations

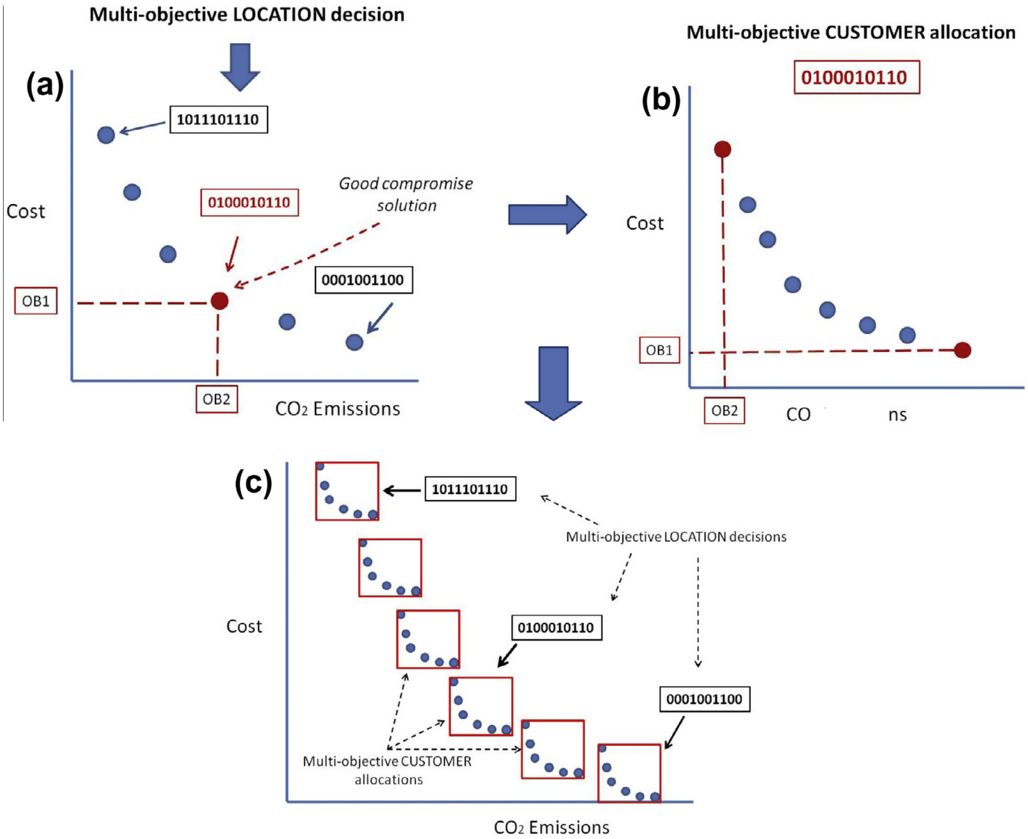

Fig. 1. For each location solution, there is a tradeoff front of allocation solutions.

separately based on cost and CO_2 emissions to assess the robustness of each candidate location solution, and (3) we identify good robust compromise solutions for *facility location*, and then explore the possibility of further multi-objective optimization for *customer allocation*.

The hierarchical and integrated nature of multi-objective optimization for facility location/allocation is illustrated in Fig. 1 which shows a typical trade-off front for facility location, with the binary strings in the boxes showing which facilities are open and which are closed (where *1* represents an open depot, and *0* a closed depot). For example the string 1011101110 indicates that facilities 1, 3, 4, 5, 7, 8, and 9 are open, with 2, 6 and 10 closed. In Fig. 1 we can also observe that a range of trade-off solutions can also be obtained for each location solution if a second multi-objective optimization process is carried out for customer allocation. The purpose of this paper is to present a method for a multi-objective optimization support toolkit from which a decision maker can explore trade-off solutions for customer allocation based on the pre-selected facility location solutions, that are of a strategic nature and not easily reversible in real life.

The integrated characteristic of solving facility location/allocation problems implies that in each iteration of the chosen solution technique, the allocation procedure allocates customers to a particular combination of open depots in order to determine an objective function. The integrated nature of the problem does not imply that we cannot separate those subproblems (location and allocation), where the facility location stage represents decision making at the *strategic level*, and the customer allocation stage gives flexibility at the *tactical stage*. In most circumstances, opening and closing facilities involves huge capital investment and thus a solution at this level we would expect to stay in place probably for decades. On the other hand, it is relatively easy to reassign customers to be supplied by different facilities, and this can be done more frequently. Using a decision support toolkit capable of exploring multi-objective optimization at the strategic and tactical levels simultaneously, provides scope for well-informed and flexible management decision making.

We were fortunate to have the opportunity to collaborate with a major UK supermarket chain during the past few years, which has given us a practical insight into the operation of a real-world distribution network. Applying this knowledge, we have generated a range of large test instances with realistic characteristics (which can be obtained from HarrisData (2011)) to test our multi-objective approach. Furthermore we use the Company's model to evaluate the various costs associated with the different warehouse and transport activities within the supply chain. Although locations, capacities and levels of demand, etc. are all randomly generated for our test instances, reasonable boundaries on these values have been observed based on Company data. However, although we are using instances based on a particular retail company in this paper, our general approach can be easily adapted for other scenarios.

We have used UK Government sources (DEFRA, 2005) to obtain correct environmental information regarding energy consumption. In addition, similar to the Company's logistic model, our data sets have two capacity constraints for each potential facility: one for the maximum number of cases that can be stored, and one for the maximum number of customers that can be served (note: this relates to space restrictions for vehicle loading). The two-objective CFLP model aims to balance the financial cost (£) and the environmental impact (kg CO_2), taking into account activities such as moving, picking, and loading the goods as well as transporting them and opening the depots required to serve the customers' needs. The environmental impact is extracted from running the logistics network in terms of CO_2 emissions from the transportation and also from the emissions generated by energy use for the day-to-day running of the depots.

4. Problem formulation

We assume that customers have a certain demand in cases (for a single product type) and associated transportation and warehousing costs for a particular depot. In our model, a single product configuration is viewed as a product category where aggregated volumes have common characteristics. Individual depots have given capacities in the number of cases they are able to store and the number of customers they are able to serve. Each customer is served directly by a single depot, and transportation costs are based on stem distances and reflect both time and distance based components. The warehouse (variable) costs reflect any costs associated with picking and loading the goods. The problem is formulated as a mixed integer programming model to determine how many facilities to open in order to satisfy all customer demand while solving both objectives simultaneously: minimize the overall financial cost and minimize the environmental impact from operating depots and transport in terms of CO_2 emissions. As mentioned above, the CFLP is divided into two sub-problems: determine which depots to open, and allocate customers to the open facilities without violating the number of cases or the number of customers' capacity constraints.

The following notation is used in the formulation of the model:

Glossary:

$V_{DC} = \{1 \ldots i\}$	set of potential depots
$V_C = \{1 \ldots j\}$	set of customers
c_{ij}	cost of attending demand from customer j to depot i consisting of transportation and depot costs
f_i	fixed cost for operating a depot i
d_j	demand of customer j
q_i	capacity (cases) of facility i, $i \in V_{DC}$
n_i	capacity (number of customers) of facility i, $i \in V_{DC}$
e_t_{ij}	CO_2 emissions from transport between depot i and customer j to satisfy customer demand d_j
e_g_i	CO_2 emissions from gas consumption for each depot i, $i \in V_{DC}$
e_e_i	CO_2 emissions from electricity consumption for each depot i, $i \in V_{DC}$

The decision variables are:

x_{ij}	equals 1 if customer j is allocated to facility i, and 0 otherwise
y_i	equals 1 if depot is chosen to operate and 0 otherwise

The following economic and environmental objective functions are considered simultaneously as part of the network design:

- *Minimizing costs*

The objective aims to find the best number of open depots with associated allocated customers that minimize total costs. It consists of the variable connection costs (transport and depots) of servicing the demand of all customers plus the fixed cost of running the open depots. The total connection cost c_{ij} from depot i to customer j as a total of $(tc_{ij} + dc_{ij})$, where tc_{ij} is the total transportation cost and dc_{ij} is the associated depot cost to supply a demand to customer j from depot i.

$$\text{Minimize} \quad \left[\sum_{i \in V_{DC}} \sum_{j \in V_C} c_{ij} x_{ij} + \sum_{i \in V_{DC}} f_i y_i \right] \tag{1}$$

- *Minimizing the CO_2 emissions from transport and depots*

The objective that aims to find the best number of open facilities that minimizes the total CO_2 emissions from transportation and energy consumption for running facilities. The first term represents the emissions from transport to attend the demand of customers by the open depots, and the second term represents the total emissions from the electricity and gas usage of operating open depots.

$$\text{Minimize} \quad \left[\sum_{i \in V_{DC}} \sum_{j \in V_C} e\_t_{ij} x_{ij} + \sum_{i \in V_{DC}} (e\_g_i + e\_e_i) y_i \right] \tag{2}$$

Both (1) and (2) are subject to following constraints:

$$\sum_{i \in V_{DC}} x_{ij} = 1, \quad j \in V_C \tag{3}$$

$$\sum_{j \in V_C} d_j x_{ij} \leqslant q_i y_i, \quad \forall i \in V_{DC} \tag{4}$$

$$\sum_{j \in V_C} x_{ij} \leqslant n_i y_i, \quad \forall i \in V_{DC} \tag{5}$$

$$x_{ij} \in \{0, 1\}, \quad i \in V_{DC}, \quad j \in V_C \tag{6}$$

$$y_i \in \{0, 1\}, \quad i \in V_{DC} \tag{7}$$

Constraints (3) and (6) ensure that each customer is attended by only one depot and the demand is satisfied by that facility. (4) and (5) ensure that the capacity constraints for demand (cases) and number of customers for the depots are not violated. Finally, (6) and (7) define decision variables as binary.

5. Multi-objective optimization

The evolutionary multi-objective algorithm SEAMO2 (Mumford, 2004; Valenzuela, 2002) (see Fig. 2) was chosen for our application in modeling large data sets. Following exploratory tests on the uncapacitated facility location problem we found that this algorithm is considerably faster than better known algorithms such as NSGA-II with small reductions in solution quality (see also Colombo and Mumford, 2005). We have compared NSGA-II (which is generally considered as the leading multi-objective evolutionary algorithm) and SEAMO2 in two tests: (1) under equal number of evaluations for each run (10,000) and (2) under equal execution time for 20 runs (200,000 ms in total).

From the experiments in test 1, we observed that SEAMO2 algorithm was much faster than NSGA-II (e.g. 53,909 ms for NSGA-II and 532 ms for SEAMO2) although the NSGA-II algorithm obtained solutions that were slightly better compared to SEAMO2. The experiments in test 2, illustrated that visually the final approximated Pareto frontier indicated very little difference. This means that for large size data sets, SEAMO2 would be able to find non-dominated solutions much more quickly

```
Begin:
Generate N random individuals
Evaluate the two objectives for each population member and store them.
Store best-so-far values for each objective
while stopping condition not satisfied do
  for each member of the population do
    This individual becomes the first parent
    Select a second parent at random
    Apply crossover to produce single offspring
    Apply single mutation to the offspring
    Evaluate each objective vector produced by the offspring
    if offspring harbors a new best-so-far Pareto component then
      a) it replaces a parent, if possible
      b) else it replaces another individual that it dominates at random
    else if offspring is a duplicate then
      it dies
    else if offspring dominates either parent then
      it replaces it
    else if offspring is neither dominated by nor dominates either parent then
      it replaces another individual that it dominates at random
    else
      otherwise it dies
    end if
  end for
end while
Print all non-dominated solutions in the final population
End
```

Fig. 2. Algorithm 1: SEAMO2.

and provide the decision-maker with an initial set of solutions, which can be explored further if desired. Nevertheless, this is essentially a proof of-concept paper that hybridizes evolutionary multi-objective optimization with Lagrangian Relaxation, and SEAMO2 could be easily substituted with NSGA-II or SPEA2, or any other evolutionary multi-objective algorithm. In addition, later on, in Section 11, we will demonstrate that SEAMO2 has found excellent trade-off solutions, sometimes even coinciding with the true Pareto-optimum frontier for the extreme points (which we find using CPLEX on some smaller problems).

5.1. SEAMO2

In a multi-objective context, fitness functions are usually based either on a count of how many contemporaries in the population are dominated by a particular individual, or alternatively on by how many contemporaries the individual is itself dominated. This approach, known as *Pareto-based selection*, was first proposed by Goldberg (Goldberg, 1989), and is favored, in one form or another, by most researchers (for example, see (Corne et al., 2000; Deb et al., 2000; Zitzler et al., 2001). In contrast, the SEAMO algorithms use a less compute-extensive approach, based on uniform random selection and do not require global fitness functions either to bias the selection of parents or to determine whether or not new offspring are inserted into the population. Instead a few simple rules determine "who shall live and who shall die" within a steady-state environment. Survival decisions are based on the outcome of simple comparisons between offspring solutions and parents (or other population members). Fig. 2 outlines the SEAMO2 algorithm.

At the beginning, the algorithm generates the initial population and evaluates all objective functions for each member of the population. Best-so-far values are recorded and the algorithm iterates through two nested loops. The *For* loop considers each member of the population in turn as a first parent for a crossover operation where a second parent is selected at random from other members of the population. A single offspring is produced as a result of crossover with a single mutation applied afterwards. The SEAMO2 algorithm is steady-state where a new offspring is considered for entry into the population that depends on a number of comparisons. The offspring replaces a current member of the population if any of "best-so-far" objectives are improved OR if the offspring dominates either of its parents OR the offspring is neither dominated by nor dominates either parent. The appropriate best-so-far scores are updated if the offspring produced better values, ensuring that best-so-far score for other objectives are not lost in the process. An additional test is carried out to check whether the offspring are duplicated and, if a match is found, then the offspring will die to preserve population diversity. The strategy of replacing a population member by dominating offspring allows the solutions to move closer to the Pareto frontier as the search progresses. The *While* loop repeats the algorithm until the stopping condition is satisfied and at the end of the procedure all the non-dominated solutions from the final population are saved to a file.

Solution encoding involves the use of simple binary strings, as described in Section 1, for which *1* represents an open depot, and *0* a closed depot. The *customer allocation procedure* for the CFLP is extremely important and ensures that capacities in terms of cases and numbers of customers are not violated. Here, we utilize a *Lagrangian Relaxation* (*LR*) technique for assigning the customers to open depots. As a result of applying our procedure, the customers are assigned according either to the minimum possible cost (*LR1* procedure in Section 7), or the minimum possible CO_2 emissions (*LR2* in Section 7), to give us two different allocation solutions to feed into the dual cost and CO_2 objectives. Fig. 3 illustrates the allocation procedure for different objectives. Our technique was adapted from several sources (Fisher, 1981, 1985; Jornsten and Nasberg, 1986;

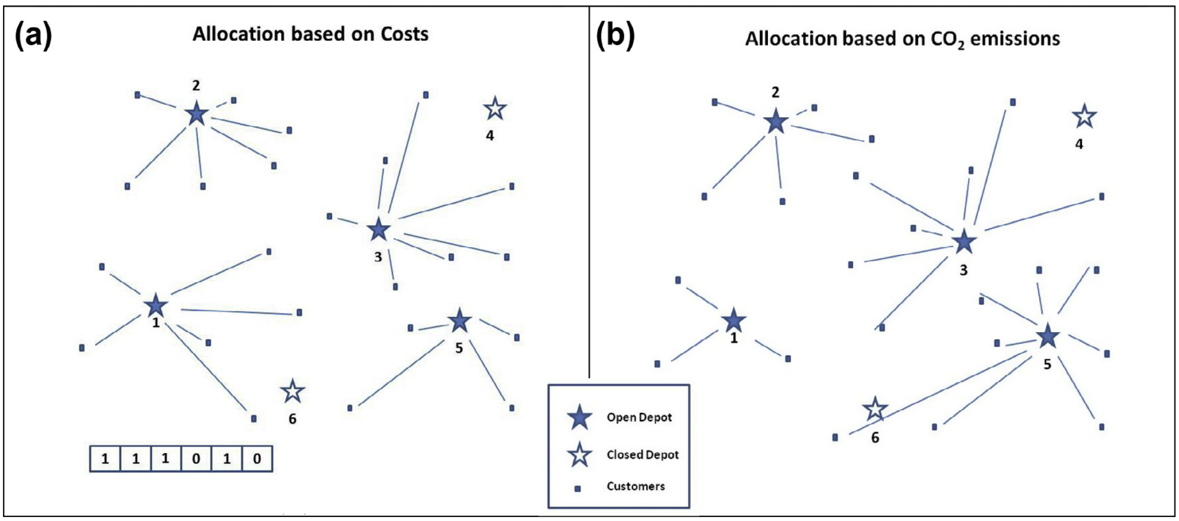

Fig. 3. Allocation of customers to depots based on cost (a) and CO_2 (b).

Ghiani et al., 2004) where it was applied to a Generalized Assignment Problem and to FLP problem where both location and allocation and with one relaxed constraint.

To summarize our multi-objective optimization approach, Fig. 4 presents an evolutionary multi-objective optimization framework where the SEAMO2 algorithm and *LR* approach are integrated together for the CFLP. The SEAMO2 algorithm determines which facilities should be open, and the LR heuristics solve the resulting Generalized Assignment Problem which allocates customers to depots. Initially, a population of solutions (binary strings) is randomly generated and for each solution we apply *LR* heuristics to allocate customers to open depots based on costs (*LR1*) for cost based optimization and CO_2 (*LR2*) emissions for CO_2 objective. Consideration needs to be taken at this stage to make sure that the randomly generated solutions with open depots are feasible, e.g. the total capacity of open depots is larger than total available demand. Following the

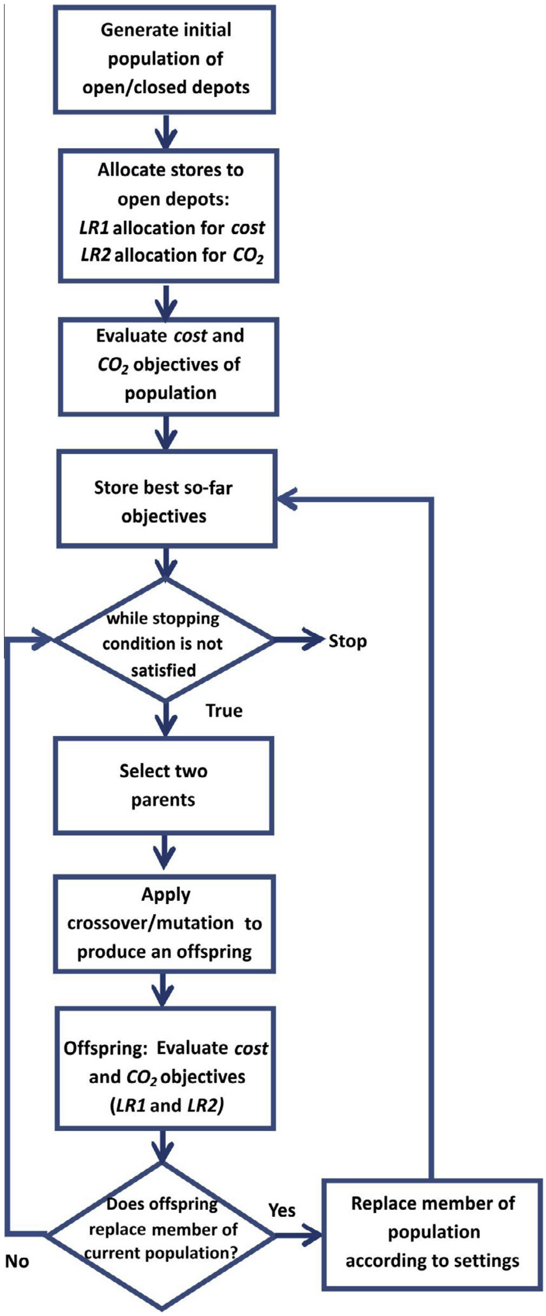

Fig. 4. Multi-objective evolutionary framework combining SEAMO2 and LR.

production of an infeasible string, one could either undertake a repair procedure to the chromosome that is infeasible or alternatively generate another random string that is feasible. The latter routine is implemented as part of the procedure.

6. Experimental setup

The main experiments deal with the location problem, for which we use SEAMO2 to determine which facilities should be open and which closed to balance the two objectives of cost and CO_2 emissions. For each facility location solution however, two different customer allocation solutions are obtained, using LR1 to optimize for cost and LR2 to optimize for CO_2 emissions. These two customer allocation solutions represent the extremes for cost and CO_2 for what can be viewed as a second level of a multi-objective hierarchy as illustrated in the Introduction by Fig. 1. In addition to the main experiments, we briefly further explore the second level of the multi-objective hierarchy for selected compromise solutions using a weighted sum *Lagrangian Relaxation* techniques, combining cost and CO_2 emissions in a single objective. The purpose of this is to offer the potential for tactical tuning to rebalance the two objectives at regular intervals of time, once the strategic facility location decision has been made.

6.1. The multi-objective facility location problem

As a result of our preliminary tests on the SEAMO2 algorithm discussed earlier, we use *uniform crossover with mutation* for the final experiments and increase the size of the population to 100 and the number of generations to 1000 to allow the algorithm to run for an adequate amount of time to find good solutions. The final non-dominated set was extracted as a result of 2 independent replicate runs for each problem instance. Initially, we performed 10 independent runs for each problem and when we analyzed final solutions, we could see that the final approximated Pareto frontier has the same results after 2 and 10 runs. Therefore we present results in this paper as a result of running the algorithm 2 independent replicate runs.

The quality of the solutions is compared to a Pareto-optimum set (based on cost or CO_2 emissions) which we obtained using CPLEX optimization software for some of the instances. We solved the allocation of the customers based on a single objective (cost or CO_2 emissions) for all possible combinations of open depots for 2000 and 4000 customers. In addition, in our previous publication (Harris et al., 2011b), we published computational times and optimum solutions located at the extreme points (min cost and min CO_2 emissions) which we solved using CPLEX. We will utilize those results as part of our discussion in the next section. CPLEX formulates the CFLP problem as a mixed integer programming problem which balances optimality and feasibility in its search using a dynamic search methodology (IBM©, 2009). The dynamic search algorithm consists of LP relaxation, branching, cuts and heuristics to find a solution. All experiments were conducted on a PC with Intel® Core™ i3 CPU, 550 @ 3.20 GHz, 3.19 GHz, 3.17 GB of RAM.

6.2. The multi-objective customer allocation problem

For the second set of experiments we start with some good compromise solutions for facility location from the previous set of experiments and explore the multi-objective trade-off front for customer allocations. We choose data instances with 2000 and 8000 customers to analyses the trade-off solutions. For each instance, we pick one good compromise solution and apply to each solution a weighted approach to the allocation of customers to depots. We use weights for costs in the range of [0,1] with step of 10 for each test.

In the multi-objective allocation of customers to the depots, we have the dilemma of balancing two objectives: minimizing overall costs and minimizing CO_2 emissions. The objectives have different units: £ and (kg CO_2), with different numerical ranges, making it difficult to choose appropriate weights to control the relative contribution of each objective to the weighted total. Therefore, we *normalize* the objectives to bring them so that each one typically produces values between 0 and 1. The formulation of the objective function is a sum of the weighted *normalized* objectives, which converts the problem into single-objective optimization problem: $z_{ij} = w_1 c^*_{ij} + w_2 e\_t^*_{ij}$.

To convert each value c_{ij} and e_t_{ij} to a single *normalized* value (c^*_{ij} or $e\_t^*_{ij}$) which is used for the composite variable z_{ij}, the following procedure was used to generate a number between 0 and 1:

- Find the highest and the lowest value of the data set.
- Calculate the *normalized* value for each store to each depot: *normValue* = (*aValue* − *minVal*)/(*maxVal* − *minVal*), where *aValue* is the corresponding value used for the optimization function, either total cost for all products from a store to the depot or the overall CO_2 emissions.

The two weights w_1 and w_2 are chosen in such way where one weight is independent and the other one is calculated by simple subtraction. Therefore, the sum of the weights is equal to 1.

As a result of the above procedure, after applying weights to the *normalized* values, a matrix of the *weighted-normalized* values is produced (z_{ij}), where each element of the matrix corresponds to the value of the allocation between each store and each depot. The matrix z_{ij} is used instead of the c_{ij} values in the LR technique to allocate customers.

7. Lagrangian Relaxation formulation for the generalized assignment problem

Fig. 4 outlines our methodology and indicates the role played by Lagrangian Relaxation (LR). LR is a general technique that can be applied to various combinatorial optimization problems where certain constraints make the problem difficult to solve. The main idea is to relax the problem by removing the difficult constraints and putting them instead into the objective function, assigning weights (called *Lagrangian Multipliers*), such that each weight represents a penalty which is added to a solution that does not satisfy the particular constraint. The more the solution violates the constraint, the higher the penalty imposed.

In the case of the Generalized Assignment Problem, two alternatives are commonly presented in the literature, either to relax the demand constraint or relax the capacity constraint (Fisher, 1981, 1985; Jornsten and Nasberg, 1986) and in this paper we relax two capacity constraints for a specific problem formulation. Given a particular subset of open facilities, ideally it would be best to assign each customer to its nearest (or lowest cost) facility. However, capacity constraints impose limitations, so that more distant or higher cost allocations may have to be applied to some customers to assure a feasible solution.

7.1. Relaxing two constraints: number of cases and number of stores

The *Lagrangian Relaxation* is applied to the allocation of stores to open depots only, i.e., the Generalized Assignment Problem. The formulation for a bi-objective problem is presented below where $k \in V_{DC\_open}$ (V_{DC_open} is a set of open depots). Formula (8) aims to minimize the total connection cost of assigning demand of store j to open depot k, whereas the formulation (9) minimizes environmental impact of transportation from store j to open depot k to deliver demand d_j. Please note that fixed costs associated with open depots that appear in the formulation of the CFLP in Eqs. (1) and (2) are not needed here. The determination of which depots to open is solved by the SEAMO2 algorithm, and the associated transportation problem is solved using the LR technique.

$$\text{Minimize} \quad \sum_{k \in V_{DC\_open}} \sum_{j \in V_C} c_{kj} x_{kj} \tag{8}$$

$$\text{Minimize} \quad \sum_{k \in V_{DC\_open}} \sum_{j \in V_C} e\_t_{kj} x_{kj} \tag{9}$$

where (8) and (9) are subject to following constraints:

$$\sum_{k \in V_{DC\_open}} x_{kj} = 1, \quad j \in V_C \tag{10}$$

$$\sum_{j \in V_C} d_j x_{kj} \leqslant q_k, \quad \forall k \in V_{DC\_open} \tag{11}$$

$$\sum_{j \in V_C} x_{kj} \leqslant n_k, \quad \forall k \in V_{DC\_open} \tag{12}$$

$$x_{kj} \in \{0,1\}, \quad k \in V_{DC\_open}, \ j \in V_C \tag{13}$$

Constraints (10) and (13) ensures that each store is assigned to one depot only and (13) defines a decision variable as binary. Constraints (11) and (12) guarantee that the capacity constraints are not violated.

In our previous work (Harris et al., 2011b), we relaxed just one capacity constraint: the number of cases. In this section we present a modified solution technique in which two capacity constraints are relaxed simultaneously: (1) the number of cases that an individual depot can handle, and (2) the number of stores that it can serve. This model is based on the requirements of the retail company that we worked with. Lagrangian multipliers are defined for the number of cases and number of stores, respectively: $\lambda_k \in R_{\geqslant 0}$ and $\varphi_k \in R_{\geqslant 0}$ for $\forall\ k \in V_{DC\_open}$. The following relaxed economic and environmental objective functions are considered simultaneously to obtain two allocation solutions at level (2) for each level (1) solution giving:

- *LR1:* Minimization by cost

$$\sum_{k \in V_{DC\_open}} \sum_{j \in V_C} c_{kj} x_{kj} + \sum_{k \in V_{DC\_open}} \lambda_k \left(\sum_{j \in V_C} d_j x_{kj} - q_k \right) + \sum_{k \in V_{DC\_open}} \varphi_k \left(\sum_{j \in V_C} x_{kj} - n_k \right) \tag{14}$$

- *LR2:* Minimization by CO_2 emissions

$$\sum_{k \in V_{DC\_open}} \sum_{j \in V_C} e\_t_{kj} x_{kj} + \sum_{k \in V_{DC\_open}} \lambda_k \left(\sum_{j \in V_C} d_j x_{kj} - q_k \right) + \sum_{k \in V_{DC\_open}} \varphi_k \left(\sum_{j \in V_C} x_{kj} - n_k \right) \tag{15}$$

- *LR1 and LR2* are both subject to following constraints:

$$\sum_{k \in V_{DC\_open}} x_{kj} = 1, \quad \forall j \in V_C \tag{16}$$

$$x_{kj} \in \{0,1\}, \quad k \in V_{DC\_open}, \ j \in V_C \tag{17}$$

Problem (14)–(17) can be decomposed into |VC| sub-problems for each relaxation: LR1 and LR2. For a given set of multipliers, $\lambda_k \in R_{\geqslant 0}$, $\varphi_k \in R_{\geqslant 0}$ the optimal lower bound of the problem (14)–(17), $LB\ (\lambda^t \varphi^t)$, at time step, t, can be found by solving the following sub-problems for each customer $j \in V_C$.

- *LR1:* Minimization by cost

$$\sum_{k \in V_{DC\_open}} (c_{kj} + d_j \lambda_k + \varphi_k) x_{kj} \tag{18}$$

- *LR2:* Minimization by CO_2 emissions

$$\sum_{k \in V_{DC\_open}} (e\_t_{kj} + d_j \lambda_k + \varphi_k) x_{kj} \tag{19}$$

- *LR1 and LR2* subject to:

$$\sum_{k \in V_{DC\_open}} x_{kj} = 1, \quad \forall j \in V_C \tag{20}$$

$$x_{kj} \in \{0,1\}, \quad k \in V_{DC\_open}, j \in V_C \tag{21}$$

- For LR1 and LR2 we set:

$$LB(\lambda^t \varphi^t) = \sum_{j \in V_C} LB^j(\lambda^t \varphi^t) - \sum_{k \in V_{DC\_open}} \lambda_k q_k - \sum_{k \in V_{DC\_open}} \varphi_k n_k \tag{22}$$

(18) and (19) are easily solved by applying a greedy algorithm to allocate each customer along lowest augmented costs, c_{kj} + $d_j\ \lambda_k + \varphi_k$ (*LR1*) or $e\_t_{kj} + d_j\ \lambda_k + \varphi_k$ (*LR2*). To provide good updating formulae for the Lagrangian multipliers, we will need upper bounds, in addition to the lower bounds in (22).

For an upper bound (*UB*) we will use a feasible solution obtained on the basis of the allocations of customers to facilities discovered in the evaluation of $LB\ (\lambda^t \varphi^t)$. However, it is likely that the allocation made for the lower bound calculation will produce some capacity violations. In order to obtain the best possible upper bounds for *LR1* and *LR2* we assert that it is better to allocate customers with high demand first, to try to ensure that individual depots have sufficient unused capacity. Therefore, the procedure sorts customers in non-increasing order of demand level (highest demand first), then works through the list, assigning customers in the same way as they were assigned to compute the *LB*, whenever possible. When capacity constraints are violated for the *LB* allocation, we iterate through the sorted list of depots, attempting to assign on the basis of the next lowest augmented cost depot, until a legal allocation is found, or the list is exhausted.

7.2. Updating the Lagrangian multipliers

For each facility at time step, t.

$$s_k^t = \sum_{j \in V_C} x_{kj}^t d_j - q_k \tag{23}$$

$$r_k^t = \sum_{j \in V_C} x_{kj}^t d_j - n_k \tag{24}$$

where x_{kj}^t is the solution of the Lagrangian Relaxation (14) or (15), subject to constraints (16) and (17), using $\lambda_k^t \in R_{\geqslant 0}$ and $\varphi_k^t \in R_{\geqslant 0}$; $\forall k \in V_{DC\_open}$ as the Lagrangian multipliers. Now set

$$\lambda_k^{t+1} = \max(0, \lambda_k^t + \beta^t s_k^t) \tag{25}$$

$$\varphi_k^{t+1} = \max(0, \varphi_k^t + \gamma^t r_k^t) \tag{26}$$

$$\beta^t = \frac{\alpha(UB - LB(\lambda^t \varphi^t))}{\sum_{k \in V_{DC\_open}} (s_k^t)^2} \tag{27}$$

$$\gamma^t = \frac{\alpha(UB - LB(\lambda^t \varphi^t))}{\sum_{k \in V_{DC\_open}} (r_k^t)^2} \tag{28}$$

where β^t and γ^t are suitable scalar coefficients where α is a constant in the interval (0,2] (Ghiani et al., 2004).

The procedure will start by seeding all the Lagrangian multipliers to zero. The value of λ_k will increase at a certain facility k, if s_k^t is positive, meaning that demand outstrips supply for that facility and decrease if a facility has a spare capacity to make associated costs more attractive.

In this section, we present a *Lagrangian Relaxation* for a single product formulation. In addition, we have developed a LR solution formulation for a multiple product environment, presented in Appendix A. Further work is needed to test a variety of settings for multi-products, to ensure that the presented technique produces efficient solutions. Testing the theoretical formulation presented in the Appendix A is outside the scope of the present paper where our contribution emphasizes the MOO integrated approach with flexibility and robustness at the allocation level.

8. Test data

There are a number of data instances available in the public domain to solve a single source single objective facility location problem (OR-Library, 2009; Avella and Boccia, 2009; Delmaire et al., 1999). Initially we analyzed those instances with an aim of adjusting them to solve a multi-objective problem formulation with environmental aspects. Unfortunately, those test instances only have demand, capacity and cost elements and there is no information regarding distances or actual locations of serving facilities and demand points (x and y coordinates) which are needed to determine environmental impact from transportation and serving facilities and apply multi-objective optimization. For that reason we generated our own random data sets (HarrisData, 2011) which are based on the aforementioned Company data, in terms of the approximate balance of demand, transport and productivity costs, with some random variation added.

We model our financial costs with both transport (distance and time elements) and depot related components based on the actual cost structure used by the retail company that we worked with. Recall, we calculate the total connection cost c_{ij} as a total of transportation and the associated depot cost to supply a demand to customer j from depot i. The cost function related to the transport component has explicit distance and time related elements where we use Euclidian distance formulae to calculate the distance between two points, and from this distance we derive the travelling time of the vehicle (the time element). We also reflect on the fact that depot costs can vary due to the geographical location of the facility; therefore each depot has its own rates for transport and warehouse components. The environmental data is generated based on the CO_2 emissions from gas and electricity per a case of demand, per week. Depending on the available capacity of the serving facilities, each of those values is multiplied by the total capacity to calculate a total energy consumption in kW h which we converted to CO_2 emissions using conversion factors from DEFRA (2005) (0.54 kg CO_2 per kW h for electricity and 0.19 kg CO_2 for gas). To compute CO_2 emissions from the transport we calculated the total distance travelled by a vehicle to satisfy a demand of a particular customer from a depot, which was multiplied by a fuel conversion factor (2.63) and then multiplied by fuel consumption (of 0.35 l per km) (DEFRA, 2005). This approach has been commonly used in the sustainable literature (Edwards et al., 2010; Harrison et al., 2010; Ubeda et al., 2011).

To ensure that feasible solutions exist for each generated instance, we produced a range of different capacities for depots which are relative to the overall demand across all depots. In this way we tested different ratio values (2, 3, 4, 5, 6, 7, 8, 9) relating the overall capacity of the total network in number of cases to the total number of customers in the network. As a result, for any given problem instance, every depot in that instance will have exactly the same capacity constraint. The fixed costs for each depot depends on the capacity of the depot and we used different ratio values (0.5, 0.75, 1.25 and 1.5) to generate instances with different features. As a result of initial experiments, we decided to use ratio values of 4 and 8 for the capacity and a fixed cost ratio of 1.25 to generate data instances. In total we generated 10 different test instances for dual objectives where financial and environmental aspects are solved simultaneously. Our instances have 10 depots and the number of customers varies thus: *2000*, *4000*, *6000*, *8000*, *10,000*. The names of the instance reflect these settings. For example, data instance *set1_10_6000_r4.0_fc1.25.txt* has 10 depots, 6000 customers, a value of 4 for the capacity ratio and the fixed cost ratio is equal to 1.25.

9. Tuning the SEAMO2 algorithm

Before running experiments on the generated data sets, SEAMO2 was tuned to its best performance on two data instances: *set1_10_2000_r4.0_fc1.25* and *set1_10_8000_r4.0_fc1.25*. We use the S metric (hypervolume measure) (Zitzler, 1999) to compare the volume of the dominated space for different types of crossover and mutation: no crossover/no mutation, one-point crossover/mutation, two-point crossover/mutation, and uniform crossover/mutation. A population size of 40 with a number of generations of 250 was used for different types of setting. In total, 8 different experiments were undertaken for tuning purposes. The final approximated Pareto frontier and the S metric were obtained from 20 independent runs for each data instance and crossover/mutation setting. As a result of undertaking one-way analysis of variance on the S metric for different sets of experiments, results indicate no statistical significance among all settings. Nevertheless, we choose *uniform crossover with mutation* to bring diversity into the population of solutions and because it appeared to have better performance compared to other settings, even though the difference was not statistically significant.

10. Testing LR technique on the allocation problem

Prior to integration of our LR into the multi-objective optimization framework, it was essential to test the technique to solve an allocation problem for the quality of the solution. It would be desirable to test the method on the benchmarking data instances available in the public domain with known solutions to compare the results. Unfortunately, we could not find allocation data which has two constraints: capacity in the number of cases and number of stores. As a result, we decided to test the Lagrangian heuristic without relaxing the number of stores constraint on the 5 medium size and 12 largest benchmark data instances of Beasley (1988) which are available in the literature for a single source FLP from the OR-library (2009). The chosen benchmarks however, are all facility location/allocation problems. Thus it was necessary to begin with solutions to the location problems, with the open facilities defined at the start, before we attempted to solve the allocation problems, and carry out the solution quality comparison for LR. Unfortunately, although optimum solutions in terms of cost are given for small and medium instances, for large problems which were used for testing there was no available cost solutions, only information on the *LB*. Also, there is no information available to identify which depots are open and which closed in the optimum solutions. To determine which depots are open and which are closed in the optimum solutions as well as their solutions, all instances were solved by us using CPLEX.

Table 2 presents the solution quality of the test results of the Beasley benchmarking instances. As can be seen from the table, for medium size instances (cap64, cap93, cap94, cap123, cap124), which range in number of facilities and number of demand points (customers) from 16×50, 25×50 and 50×50, our LR has found optimal solutions. For larger instances (capa1–capc4) with 100 potential facilities and 1000 customers, our approach has found very good quality solutions for all instances. As a result of those experiments, we proceeded further to integrate our LR into multi-objective optimization framework. Further analysis is presented in Section 11 to demonstrate the quality of the LR technique presented in the Section 7 where both constraints are relaxed simultaneously: number of cases and the number of stores that a depot can serve.

11. Results

11.1. Facility location

For each problem instance, we obtain the final Pareto frontier for the location problem, as a result of 2 independent runs of SEAMO2. Figs. 5 and 6 illustrate the approximated Pareto frontier for three data instances with 2000, 6000 and 8000 customers where we have different allocations of customers to the serving depots for each objective function: one objective based on costs and the other based on CO_2 emissions. Fig. 5 also illustrates Pareto-optimum solutions which were found by solving each objective function separately using CPLEX for different combinations of open facilities, where an objective functions uses an allocation procedure which is related to the main objective function (e.g. cost for the cost objective function). We can see that our solutions and solutions found by CPLEX are almost identical, giving us confidence in our technique to produce excellent quality solutions. Table 4 presents further comparison of the results between two approaches.

Due to the characteristics of the data sets (constraints on number of customers and number of cases), we can see from the figures that the solutions are not evenly spread across the approximate Pareto front because of the constrained solution space which is reflected in a small number of feasible solutions. A similar pattern can be seen for other instances. A good compromise solution is also highlighted in Fig. 5, where the costs are still low before the Pareto frontier gets steeper towards high costs and lower emissions. In Fig. 6(b), we could identify three good compromise solutions which are located in the middle of the Pareto frontier. Table 3 presents the approximate Pareto frontiers for all our instances, including related information on open depots for each solution. Each solution from the non-dominated set has a different combination of open

Table 2
Data sets from Beasley (1988), available from OR-library (2009) to solve a single source allocation problem.

Data instance 1	Beasley solution 2	CPLEX solution (0.01% tolerance) 3	LR[c] 4	% Diff (2) vs (4) 5
cap64	1,053,197.44[a]	1,053,197.44	1,053,197.44	0
cap93	900,760.11[a]	900,760.11	900,760.11	0
cap94	950,608.42[a]	950,608.42	950,608.43	0
cap123	898,266.08[a]	898,266.07	898,266.08	0
cap124	950,608.43[a]	950,608.42	950,608.43	0
capa1	19,240,822.45[b]	19,241,057.80	19,242,654.81	0.009523
capa2	18,438,046.54[b]	18,438,329.78	18,439,832.15	0.009684
capa3	17,765,201.95[b]	17,765,201.95	17,765,201.95	0
capa4	17,160,439.01[b]	17,160,815.54	17,161,398.06	0.005589
capb1	13,656,379.58[b]	13,657,482.15	13,659,675.22	0.024133
capb2	13,361,927.45[b]	13,363,068.68	13,364,719.37	0.020895
capb3	13,198,556.43[b]	13,199,420.27	13,208,051.42	0.071940
capb4	13,082,516.50[b]	13,083,451.13	13,086,320.44	0.029077
capc1	11,646,596.97[b]	11,647,531.06	11,649,860.05	0.028017
capc2	11,570,340.29[b]	11,570,437.68	11,570,437.68	0.000842
capc3	11,518,743.74[b]	11,519,413.38	11,520,815.85	0.017989
capc4	11,505,767.39[b]	11,505,861.86	11,505,861.86	0.000821

[a] Optimum solution.
[b] Published LB by Beasley.
[c] Final LR solution has fixed costs added for open facilities.

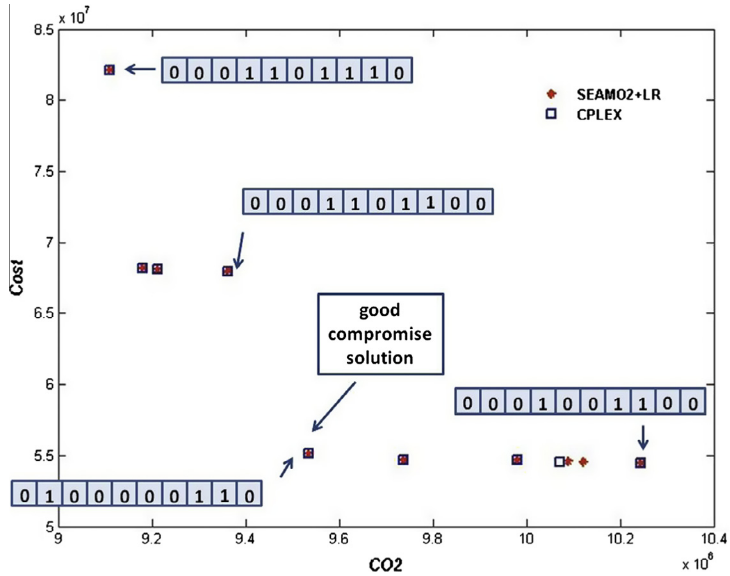

Fig. 5. Pareto frontier (SEAMO2 + LR) for instance set1_10_2000_r4.0_fc1.25.txt.

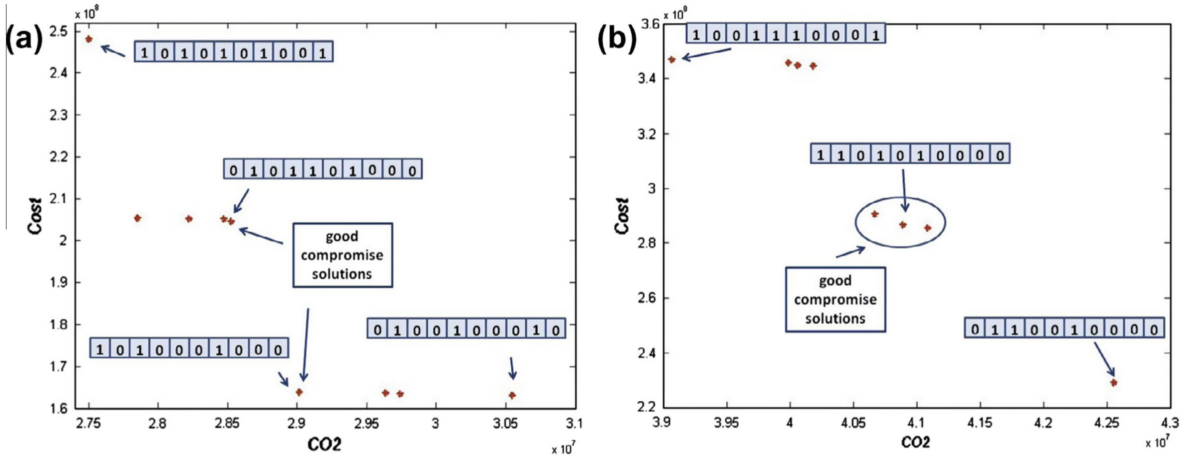

Fig. 6. Pareto frontier (SEAMO2 + LR) for instances set1_10_6000_r4.0_fc1.25.txt (a) and set1_10_8000_r4.0_fc1.25.txt (b).

depots. The results presented in the table illustrate that the lowest cost solution produces the highest CO_2 emissions with less facilities open and the higher number of open facilities produces lowest emissions at much higher cost. Table 3 also presents information regarding computational times for our MOEA framework as a result of the sum of the times for the 2 independent runs.

Table 4 compares the approximated Pareto frontier obtained using our multi-objective optimization framework with the optimum Pareto frontier obtained by CPLEX, where we solved the allocation problem twice for each open/closed depot solution produced by our multi-objective optimization: once based on costs and once based on CO_2 emissions. We also compare run times for SEAMO2 + LR to CPLEX. To help combat the "out of memory" errors and extremely slow run times for CPLEX, we applied different tolerance levels depending on the instance size. On the other hand, SEAMO2 + LR had no problems finding excellent quality solutions for all instances in a reasonable amount of time. Table 4 presents solutions for 2000 and 4000 customers only because we could not solve other instances using CPLEX due to the size of the problems and out of memory error, even when applying very generous tolerance levels. For examples, in the instance 2000_r4.0, solution number (1) has following open depots, *0001001100*. As a result of using the SEAMO2 + LR approach, where allocation is based on costs only (LR1), the result is equal to 54,476,666.18 whereas the CPLEX solution for the same combination of open depots and allocation based on costs is 54,476,666.07 which is almost identical. The differences in results for both techniques is 0% in the

Table 3
Non-dominated solutions and their location decision for our multi-objective optimization framework (SEAMO2 + LR).

Data instance	Solut. num.	Depots open	Num. of open depots	MOEA framework: SEAMO2 + LR		Comput. times (s) for 2 runs
				Cost solution (LR1)	CO_2 solution (LR2)	
2000_4	1	0 0 0 1 0 0 1 1 0 0	3	54,476,666.18	10,245,296.02	525.30
	2	0 0 0 1 1 0 1 0 0 0	3	54,543,121.21	10,120,643.77	
	3	0 0 0 1 0 0 0 1 0 1	3	54,582,187.15	10,089,699.03	
	4	0 0 0 1 0 0 0 1 1 0	3	54,662,867.62	9,980,595.66	
	5	0 0 0 1 1 0 0 0 1 0	3	54,669,493.25	9,737,911.86	
	6	0 1 0 0 0 0 0 1 1 0	3	55,117,117.89	9,535,213.46	
	7	0 0 0 1 1 0 1 1 0 0	4	67,964,106.89	9,361,885.46	
	8	0 0 0 1 1 0 0 1 0 1	4	68,071,418.31	9,212,343.08	
	9	0 0 0 1 1 0 0 1 1 0	4	68,196,328.12	9,180,819.04	
	10	0 0 0 1 1 0 1 1 1 0	5	82,133,809.34	9,110,719.54	
2000_8	1	0 0 0 0 1 0 0 0 1 0	2	71,012,035.53	11,882,881.70	51.89
	2	0 0 0 1 1 0 1 0 0 0	3	98,768,829.52	11,832,250.13	
	3	0 0 0 1 0 0 0 1 1 0	3	98,930,795.62	11,795,481.15	
	4	0 0 0 1 1 0 0 0 1 0	3	98,937,421.25	11,552,797.35	
	5	0 1 0 0 0 0 0 1 1 0	3	99,374,607.12	11,350,098.95	
4000_4	1	0 0 0 1 0 0 0 0 1 1	3	110,303,082.99	19,905,487.97	1881.61
	2	0 1 0 1 0 0 0 0 1 1	4	138,469,639.83	19,746,015.67	
	3	1 1 0 0 0 0 0 0 1 1	4	138,756,872.21	19,640,782.76	
	4	1 0 0 1 0 0 0 0 1 1	4	138,979,234.53	19,635,618.93	
	5	0 0 0 1 1 0 1 0 1 0	4	139,462,224.55	19,280,459.13	
	6	0 1 0 1 1 0 1 0 1 0	5	167,661,688.64	19,114,557.23	
4000_8	1	0 1 0 0 0 0 0 0 0 1	2	142,865,836.50	25,254,151.79	52.49
	2	0 1 0 0 0 0 0 1 0 0	2	143,031,976.37	25,041,064.94	
	3	0 0 0 1 0 0 1 0 0 0	2	144,413,887.67	24,601,882.11	
	4	0 0 0 1 0 0 0 0 1 1	3	199,746,328.83	23,551,122.74	
6000_4	1	0 1 0 0 1 0 0 0 1 0	3	163,089,682.86	30,545,269.91	1859.84
	2	1 1 0 0 1 0 0 0 0 0	3	163,393,425.27	29,741,572.60	
	3	1 1 0 0 0 0 1 0 0 0	3	163,675,478.82	29,637,868.06	
	4	1 0 1 0 0 0 1 0 0 0	3	163,912,900.43	29,014,063.12	
	5	0 1 0 1 1 0 1 0 0 0	4	204,555,315.01	28,527,877.85	
	6	1 1 0 0 1 0 1 0 0 0	4	205,142,703.92	28,474,622.99	
	7	0 0 1 1 1 0 1 0 0 0	4	205,211,886.15	28,222,267.98	
	8	1 0 1 0 1 0 1 0 0 0	4	205,379,806.30	27,850,818.04	
	9	1 0 1 0 1 0 1 0 0 1	5	248,079,102.86	27,502,122.32	
6000_8	1	0 1 0 0 1 0 0 0 0 0	2	211,113,103.48	36,848,288.55	136.41
	2	1 0 0 0 0 0 1 0 0 0	2	214,541,174.15	36,374,425.08	
	3	0 1 0 1 1 0 0 0 0 0	3	296,345,373.14	35,278,512.54	
	4	1 1 0 0 1 0 0 0 0 0	3	296,924,129.45	35,184,248.66	
	5	1 1 0 0 0 0 1 0 0 0	3	297,726,788.82	35,133,669.94	
	6	1 0 1 0 0 0 1 0 0 0	3	297,963,891.20	34,509,865.00	
8000_4	1	0 1 1 0 0 1 0 0 0 0	3	229,084,156.77	42,555,198.69	3170.25
	2	0 1 1 1 0 1 0 0 0 0	4	285,584,640.10	41,089,365.28	
	3	1 1 0 1 0 1 0 0 0 0	4	286,628,670.67	40,893,973.84	
	4	0 0 0 1 1 1 0 0 0 1	4	290,474,655.33	40,669,876.35	
	5	0 1 1 1 1 1 0 0 0 0	5	344,677,476.38	40,182,455.75	
	6	1 1 0 1 0 1 0 0 0 1	5	344,799,740.03	40,064,284.46	
	7	1 1 0 1 1 1 0 0 0 0	5	345,721,506.95	39,987,064.31	
	8	1 0 0 1 1 1 0 0 0 1	5	346,842,220.41	39,069,000.26	
8000_8	1	0 1 0 1 0 0 0 0 0 0	2	295,814,178.47	54,192,971.19	198.16
	2	0 0 0 1 0 0 0 0 0 1	2	298,587,738.20	52,501,456.78	
	3	0 0 0 0 0 1 0 0 0 1	2	301,036,394.73	51,762,156.41	
	4	0 1 1 0 0 1 0 0 0 0	3	410,353,081.04	49,987,201.17	
10000_4	1	0 0 0 0 0 1 1 0 1 0	3	277,887,312.30	51,204,592.17	3118.37
	2	1 0 0 0 0 1 1 0 0 0	3	278,014,186.52	50,070,358.46	
	3	1 0 0 0 0 0 1 0 0 1	3	280,034,198.13	49,393,399.29	
	4	0 0 0 1 0 1 1 1 0 0	4	348,882,397.44	47,760,427.70	
	5	1 0 0 0 0 1 1 1 0 0	4	349,262,497.09	47,741,318.20	
	6	1 1 0 0 0 1 1 0 0 0	4	349,713,573.22	47,484,274.11	
	7	0 0 0 1 0 1 1 1 1 0	5	420,200,675.86	47,414,615.01	
	8	0 0 0 0 1 1 1 1 1 0	5	420,945,294.90	47,058,336.95	
	9	1 1 0 0 1 1 1 0 0 0	5	421,969,321.40	46,848,513.35	

10000_8	1	0	0	1	0	0	0	0	0	1	0	2	359,872,980.28	63,283,331.87	108.41
	2	1	0	0	0	0	1	0	0	0	0	2	360,325,667.67	59,978,160.92	
	3	1	0	0	0	0	1	1	0	0	0	3	503,642,392.32	59,276,435.88	
	4	1	0	0	0	0	0	1	0	0	1	3	505,665,142.53	58,644,745.65	

Table 4
Approximated Pareto frontier (SEAMO2 + LR) and Pareto-optimum set (CPLEX, based on costs or CO_2 emissions only) for instances with 2000 and 4000 customers.

Data instance	Solut. num.	SEAMO2 + LR		CPLEX		% Diff cost	% Diff CO_2
		Cost solution (LR1)	CO_2 solution (LR2)	Optimization by cost Cost solution	Optimization by CO_2 CO_2 solution		
2000_r4.0	1	54,476,666.18	10,245,296.02	54,476,666.07	10,245,237.42	0.00	0.00
	2	54,543,121.21	10,120,643.77	54,529,812.92	10,071,232.68	0.02	0.49
	3	54,582,187.15	10,089,699.03	–	–	–	–
	4	54,662,867.62	9,980,595.66	54,662,867.25	9,980,595.66	0.00	0.00
	5	54,669,493.25	9,737,911.86	54,669,493.28	9,737,911.86	0.00	0.00
	6[a]	55,117,117.89	9,535,213.46	55,117,943.18	9,535,213.46	0.00	0.00
	7	67,964,106.89	9,361,885.46	67,964,106.88	9,361,885.46	0.00	0.00
	8	68,071,418.31	9,212,343.08	68,071,418.04	9,212,343.08	0.00	0.00
	9	68,196,328.12	9,180,819.04	68,196,327.84	9,180,819.04	0.00	0.00
	10	82,133,809.34	9,110,719.54	82,133,809.10	9,110,719.54	0.00	0.00
2000_r8.0	1	71,012,035.53	11,882,881.70	71,012,035.52	11,882,881.70	0.00	0.00
	2	98,768,829.52	11,832,250.13	98,768,829.69	11,832,250.13	0.00	0.00
	3	98,930,795.62	11,795,481.15	98,930,795.25	11,795,481.15	0.00	0.00
	4	98,937,421.25	11,552,797.35	98,937,421.28	11,552,797.35	0.00	0.00
	5	99,374,607.12	11,350,098.95	99,374,606.89	11,350,098.95	0.00	0.00
4000_r4.0	1	110,303,082.99	19,905,487.97	110,274,829.13	19,896,917.70	0.03	0.04
	2	138,469,639.83	19,746,015.67	138,457,407.11	19,728,631.76	0.01	0.09
	3	138,756,872.21	19,640,782.76	138,748,105.78	19,640,782.76	0.01	0.00
	4	138,979,234.53	19,635,618.93	138,979,234.60	19,635,618.93	0.00	0.00
	5	139,462,224.55	19,280,459.13	139,462,224.92	19,280,459.13	0.00	0.00
	6[a]	167,661,688.64	19,114,557.23	167,661,689.01	19,114,557.23	0.00	0.00
4000_r8.0	1	142,865,836.50	25,254,151.79	142,865,836.51	25,254,151.79	0.00	0.00
	2	143,031,976.37	25,041,064.94	143,031,976.56	25,041,064.94	0.00	0.00
	3[a]	144,413,887.67	24,601,882.11	144,413,888.14	24,601,882.11	0.00	0.00
	4	199,746,328.83	23,551,122.74	199,746,328.97	23,551,122.74	0.00	0.00

[a] Small difference in the cost solution between CPLEX and SEAMO2 + LR due to the tolerance levels associated with both solutions.

majority of the solutions (20 out of 24 solutions), and for the other 4 solutions the difference is negligible. Therefore we can conclude that the SEAMO2 + LR approach produces excellent quality solutions.

The computational times of our framework vary depending on the instance and in some cases it takes only a fraction of CPLEX running time (e.g. around 4% from CPLEX running times) and for other data instances it takes much longer to find a set of solutions compared to finding one solution by CPLEX. However we are satisfied that our MOEA approach provides a set of excellent non-dominated solutions with practical application for large real-world problem instances.

Table 5 records information on solutions for the other data instances (6000, 8000 and 10,000 customers) which were obtained using CPLEX with different tolerance level settings. We compare solutions to the best known cost and CO_2 solution from SEAMO2 + LR framework. We do not have information regarding the cost solution for all instances to use them in our analysis due to CPLEX error: out of memory. As can be seen from the table, SEAMO2 + LR have found excellent quality solutions, where some of them are identical to the CPLEX solutions. This gives us an additional confidence in our approach in addition to the discussion earlier for 2000 and 4000 customers.

In addition to the above experiments, we created new data sets with 50, 100 and 200 available facilities and up to 4000 customers with different demand ratios of 3 and 6 and different settings for environmental impact in the facilities. The results of those experiments and experiments discussed in this paper for large data sets, demonstrated that our multi-objective optimization framework, consisting of the evolutionary SEAMO2 algorithm with Lagrangian Relaxation for allocating customers to the open facilities is capable to find excellent nearly "optimum" Pareto frontier solutions quickly to allow the decision maker to evaluate them for more informed decision making.

11.2. Multi-objective customer allocation

In Section 11.1, we identified a non-dominated set of solutions at the level 1 decision making stage, where the algorithms decides on which depots to be open for each data instance. In this section, we apply a weight based approach to the

Table 5
Cost and CO_2 solution comparison between SEAMO2 + LR and CPLEX solution for instances with 6000, 8000 and 10,000 customers.

Data instance	Cost solution			CO_2 solution		
	SEAMO2 + LR (best solution)	CPLEX optimization by *cost*	% Diff	SEAMO2 + LR (best solution)	CPLEX optimization by CO_2	% Diff
6000_r4.0	163,089,682.86	163,085,397.79[b]	0.0026	27,502,122.32	27,502,122.32[a]	0.00
6000_r8.0	211,113,103.48	–[c]	–	34,509,865.00	34,509,852.20[a]	0.00
8000_r4.0	229,084,156.77	–[c]	–	39,069,000.26	39,069,000.26[a]	0.00
8000_r8.0	295,814,178.47	295,814,178.47[b]	0.0000	49,987,201.17	49,987,201.17[a]	0.00
10000_r4.0	277,887,312.30	–[c]	–	46,848,513.35	46,848,513.35[a]	0.00
10000_r8.0	359,872,980.28	–[c]	–	58,644,745.65	58,644,745.65[a]	0.00

[a] Relative MIP gap tolerance: 1e−04 (by default).
[b] Relative MIP gap tolerance: 0.04.
[c] CPLEX error: out of memory.

normalized objective values to allocate customers to depots for selected good compromise solutions from some data instances. In practice, this would represent tactical choices a company could make to regularly tune customer allocation, once the strategic facility location decision had been taken. For these experiments we choose two instances with 2000 and 8000 customers. For the data instance with 2000 customers, a good compromise solution has three open depots: *0,1,0,0,0,0,0,1,1,0* and for 8000 customers, a good compromise solution has four depots open: *1,1,0,1,0,1,0,0,0,0*. Fig. 7 visualizes the approximate Pareto frontier (for 2000 customers) when different weights are applied for costs and for CO_2 emissions in the customer allocation to open depots as presented in the Table 5. As we can see from the figure, the frontier is evenly spread covering extreme points of the single objective optimization. This allows the decision maker to see straight away some good compromise solutions for allocation. For example, solutions circled in the figure have relatively low CO_2 emissions just before the curve steepens towards higher costs.

Table 6 displays the results of all non-dominated solutions of the allocation of customers to open depots based on different weights for cost ($\boldsymbol{w_c}$) and distance ($\boldsymbol{w_{CO2}}$) for 2000 and 8000 customers. As can be seen from the table, for the instance with 2000 customers, we do not present the results for $w_c = 0.2$, because that solution is dominated by the solution with the weight, $w_c = 0.3$. The results presented in the table are compared to the allocation purely based on costs alone. Weighting both objectives allows us to generate trade-off solutions and understand the relationship between cost and CO_2 emissions to effectively minimize environmental impact from the transport related emissions without a detrimental impact on the financial objective. For example, balancing cost and CO_2 objectives with weights $\boldsymbol{w_c} = 0.6$ and $\boldsymbol{w_{CO2}} = 0.4$ allow us to reduce total CO_2 emissions by around 62,155.75 units of CO_2 by increasing cost by only 14,827.01 units a week compared to the optimization based only on costs. This equates to around 3,232,099 units of CO_2 a year (52 weeks), which contributes to a significant amount of CO_2 emissions. If the decision maker only considers the percentage increase, the reduction of 0.65% in CO_2 emissions appears insignificant compared to the absolute amount of emissions saved of 62,155.75 per week.

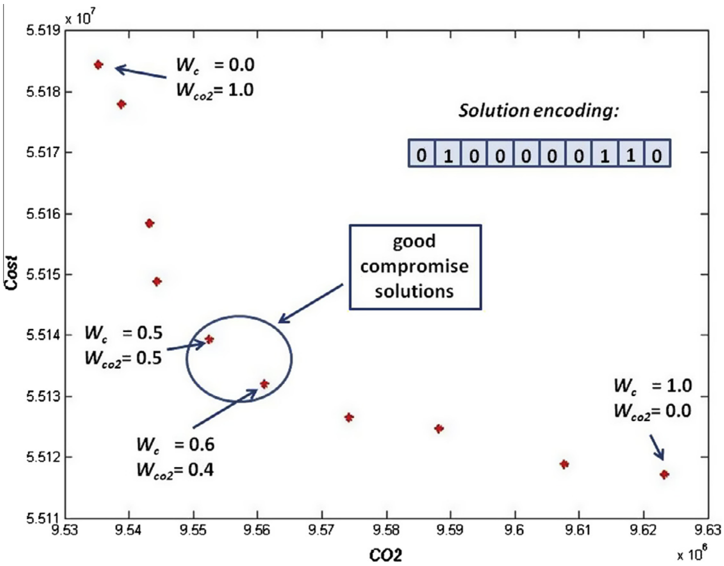

Fig. 7. Pareto frontier (SEAMO2 + LR) for the allocation with different weights to the open depots (0,1,0,0,0,0,0,1,1,0), instance set1_10_2000_r4.0_fc1.25.txt.

Table 6
Results of the optimization with different weights (SEAMO2 + LR) for $\boldsymbol{w_c}$ and $\boldsymbol{w_{CO2}}$ which compared to the allocation purely based on costs ($\boldsymbol{w_c}$ = 1, $\boldsymbol{w_{CO2}}$ = 0).

Data instance	w_c	w_{CO2}	Cost	CO_2	Difference in cost	% Diff for cost	Difference in CO_2	% Diff for CO_2
2000_r4.0	0	1	55,184,400.75	9,535,213.46	67,282.86	0.12	−87,990.17	0.91
	0.1	0.9	55,177,947.25	9,538,810.12	60,829.36	0.11	−84,393.51	0.88
	0.3	0.7	55,158,362.61	9,543,192.58	41,244.72	0.07	−80,011.06	0.83
	0.4	0.6	55,148,812.13	9,544,332.25	31,694.24	0.06	−78,871.39	0.82
	0.5	0.5	55,139,312.22	9,552,544.28	22,194.33	0.04	−70,659.36	0.73
	0.6	0.4	55,131,944.90	9,561,047.88	14,827.01	0.03	−62,155.75	0.65
	0.7	0.3	55,126,457.52	9,574,200.92	9,339.63	0.02	−49,002.72	0.51
	0.8	0.2	55,124,722.93	9,588,204.53	7,605.04	0.01	−34,999.10	0.36
	0.9	0.1	55,118,817.27	9,607,705.00	1,699.38	0.00	−15,498.64	0.16
	1	0	55,117,117.89	9,623,203.64	0.00	0.00	0.00	0.00
8000_r4.0	0	1	286,968,990.29	40,893,973.84	340,319.62	0.12	−328,588.53	0.80
	0.1	0.9	286,917,519.06	40,896,591.97	288,848.39	0.10	−325,970.40	0.79
	0.2	0.8	286,844,366.61	40,908,816.07	215,695.94	0.08	−313,746.30	0.76
	0.3	0.7	286,805,355.46	40,920,188.05	176,684.79	0.06	−302,374.32	0.73
	0.4	0.6	286,756,660.30	40,943,331.70	127,989.63	0.04	−279,230.67	0.68
	0.5	0.5	286,723,259.95	40,967,199.77	94,589.28	0.03	−255,362.61	0.62
	0.6	0.4	286,689,598.10	41,003,461.38	60,927.43	0.02	−219,100.99	0.53
	0.7	0.3	286,665,201.01	41,044,013.86	36,530.34	0.01	−178,548.52	0.43
	0.8	0.2	286,645,891.00	41,094,346.73	17,220.33	0.01	−128,215.64	0.31
	0.9	0.1	286,632,157.16	41,161,660.93	3,486.49	0.00	−60,901.44	0.15
	1	0	286,628,670.67	41,222,562.37	0.00	0.00	0.00	0.00

Therefore, looking at only % increase/decrease in the trade-off solutions may hide valuable information on the impact on the overall CO_2 emissions. Those results allow us to explore the non-dominated set of solutions and determine the impacts of the different weight settings to associated allocation.

12. Conclusion

This paper presents a new way of evaluating a multi-objective CFLP problem for large instances that considers robustness of solutions at the allocation level where financial costs and CO_2 emissions are solved simultaneously to provide a decision maker with a set of trade-off solutions. The framework utilizes an evolutionary algorithm (SEAMO2) to determine which facilities to open and a Lagrangian Relaxation technique where two constraints are relaxed to establish the best allocation of customers to serving facilities. In addition, we apply a weighted sum approach to selected good robust compromise solutions for some instances to explore a multi-objective allocation procedure which also produces good quality trade-off solutions, thus providing potential to rebalance cost and CO_2 objectives on a regular tactical basis once the strategic decisions for facility location have been made. The problem is formulated as a mixed integer programming model and uses aggregated data where (i) capacities of depots and (ii) average fixed cost for operating depot and energy consumption are known and depended on the actual capacity of the facilities. This is a valid assumption for MIP modeling, based on a number of well-known academic papers; for example, the study conducted by Beasley (1988) (the data of which is used to test our LR), considers the same formulation in relation to fixed costs. The study conducted by Holmberg et al. (1999) is another example of research where an exact algorithm was developed for CFLP.

An important contribution of the paper to the literature is that, to the best of the authors' knowledge, this is the first study that models the multi-objective facility location problem under the consideration of *flexibility* at the tactical customer allocation level; the *robustness* of the solutions over period of time with the focus on large data instances where financial and CO_2 emissions objectives are modeled simultaneously. We present the hybridization of an evolutionary multi-objective optimization algorithm through Lagrangian Relaxation as a proof of concept of an application on data instances that utilize cost and CO_2 characteristics of a 'real-world' retail secondary distribution network. Compared to recently published work, our paper presents the following advances: 1) we present and implemented a new LR model embedded into evolutionary algorithm that incorporates multi-objective allocation into every location decision where the optimization of the allocation is based on costs and CO_2 emissions; 2) we compare the results to the single objective optimum solution and to the Pareto-optimum set obtained using CPLEX software when possible. We can see from the results, discussed in the Section 11 that computational times are very good for our optimization framework and our approach finds the Pareto-optimum set for some of the instances. These trade-off results provide potentially very useful information for industry and make it possible for a decision maker to consider the balance between cost and environmental impact, and possibly find excellent solutions that would be missed using more traditional techniques; 3) we presented a solution formulation for a multiple product environment in Appendix A.

For future research, a number of interesting extensions could be investigated where the impact of uncertainty can be considered as part of developing solution techniques for stochastic models. Stochastic models for facility location are comprehensively discussed by Snyder (2006) where some models are developed under a probabilistic approach and need specifically designed algorithms or stochastic programming techniques. In relation to developing meta-heuristic techniques, Snyder (2006) mentioned that "Meta-heuristics have been applied successfully to deterministic location problems, but few, if any, have been developed for their stochastic and robust counterparts." Designing specialized algorithms for stochastic environments is a major piece of research on its own and by simply considering a sensitivity analysis in the paper would not address complex issues with regards to uncertainty. Other research could consider multi-period together with a multi-product environment and different storage requirements. The present research is a preliminary step towards a sustainable multi-objective methodology and recognizes that a range of economic and environmental variables and constraints are needed for modeling large scale secondary distribution networks.

Acknowledgement

The research presented in this paper was supported by The Engineering and Physical Sciences Research Council (EPSRC), 'Green Logistics', grant number EP/D043328/1.

Appendix A

A.1. Lagrangian relaxation formulation for multiple product formulation

The solution formulation presented in this appendix is based on the LR formulation presented in Section 7. Let $\lambda_k^p \in R_{\geqslant 0}$ and $\varphi_k^p \in R_{\geqslant 0}$ for $\forall k \in V_{DC\_open}$.

- *LR1:* Minimization by cost

$$\sum_{k\in V_{DC\_open}}\sum_{j\in V_C}\sum_{p\in V_p} c_{kj}^p x_{kj} + \sum_{k\in V_{DC\_open}}\sum_{p\in V_p}\lambda_k^p\left(\sum_{j\in V_C} d_j^p x_{kj} - q_k^p\right) + \sum_{k\in V_{DC\_open}}\sum_{p\in V_p}\varphi_k^p\left(\sum_{j\in V_C} x_{kj} - n_k^p\right)$$

- *LR2:* Minimization by CO_2 emissions

$$\sum_{k\in V_{DC\_open}}\sum_{j\in V_C}\sum_{p\in V_p} e\_t_{kj}^p x_{kj} + \sum_{k\in V_{DC\_open}}\sum_{p\in V_p}\lambda_k^p\left(\sum_{j\in V_C} d_j^p x_{kj} - q_k^p\right) + \sum_{k\in V_{DC\_open}}\sum_{p\in V_p}\varphi_k^p\left(\sum_{j\in V_C} x_{kj} - n_k^p\right)$$

- *LR1 and LR2* are both subject to following constraints:

$$\sum_{k\in V_{DC\_open}} x_{kj} = 1, \quad \forall j \in V_C$$

$$x_{kj} \in \{0, 1\}, \quad k \in V_{DC\_open}, \; j \in V_C$$

To obtain LB, the relaxed problem simply solved by applying a greedy algorithm to allocate each customer along the lowest cost. To provide a good updating formula for the Lagrangian multipliers, we will need an upper bound, in addition to the lower bound. For an UB we will use a feasible solution obtained on the basis of the allocations of customers to facilities discovered in the evaluation of *LB*. However, it is likely that the allocation made for the lower bound calculation will produce some capacity violations, especially with a multi-product problem, such as we have here. In order to obtain the best possible upper bound (i.e., with the lowest cost), we need to establish a good way of reallocating customers when facilities are over-subscribed.

For a single product problem, the customers were sorted in non-increasing order according to their demand, and then assigned to facilities based on the lower bound assignment without violating capacity constraints. When capacity constraints are violated for *LB* assignment, the customer is assigned to the next lowest augmented cost, etc. If all facilities are overcapacity then the customer is allocated to the lowest true (non-augmented cost). However, the current model formulation has multiple products where each customer will have a certain demand for each product type, and each depot with have a maximum capacity for each product type. Therefore, different settings were devised to ensure that allocating the customers produces feasible solutions when calculating *UB* assignment. The following settings could be used for finding feasible *UB* assignment:

1. {*Sorting using highest demand*} The highest demand across all product types per customer is used to sort customers in non-increasing order, where the customer with highest value is assigned first.
2. {*Sorting using highest demand and the depot load ratio*} The depot load ratio per product type is defined as the ratio of total demand of all customers per product type and total capacity of all depots for that product type. The highest demand across all product types per customer is multiplied by the load ratio of that particular product. This is used to sort customers in non-increasing order, where the customer with highest value is assigned first for *UB* assignment.

3. {*Sorting using highest fraction of normalized demand across all products*} The highest fraction of normalized demand across all product types for each customer is chosen for sorting customers in non-increasing order. The customers with highest value are assigned first.
4. {*Sorting using normalized demand and the depot load ratio*} The highest fraction of normalized demand across all product types per customer is multiplied by the load ratio of that particular product. This value is used to sort customers in non-increasing order, where the customer with highest value is assigned first for *UB* assignment.

References

Altiparmak, F., Gen, M., Lin, L., Paksoy, T., 2006. A genetic algorithm approach for multi objective optimization of supply chain networks. Comput. Ind. Eng. 51 (1), 196–215.
Avella, P., Boccia, M., 2009. A cutting plane algorithm for the capacitated facility location problem. Comput. Optim. Appl. 43 (1), 39–65.
Avella, P., Boccia, M., Sforza, A., Vasil'ev, I., 2009. An effective heuristic for large-scale capacitated facility location problems. J. Heuristics 15 (6), 597–615.
Bachlaus, M., Pandey, M.K., Mahajan, C., Shankar, R., Tiwari, M.K., 2008. Designing an integrated multi-echelon agile supply chain network: a hybrid taguchi-particle swarm optimization approach. J. Intell. Manuf. 19 (6), 747–761.
Bäck, T., 1996. Evolutionary Algorithms in Theory and Practice. Oxford University Press, New York, New York.
Barcelo, J., Casanovas, J., 1984. A heuristic Lagrangian algorithm for the capacitated plant location problem. Eur. J. Oper. Res. 15 (2), 212–226.
Beamon, B.M., 1999. Designing the green supply chain. Logist. Inform. Manage. 12 (4), 332–342.
Beasley, J.E., 1988. An algorithm for solving large capacitated warehouse location problems. Eur. J. Oper. Res. 33 (3), 314–325.
Bhattacharya, R., Bandyopadhyay, S., 2010. Solving conflicting bi-objective facility location problem by NSGA II evolutionary algorithm. Int. J. Adv. Manuf. Technol. 51 (1–4), 397–414.
Bloemhof, J.M., Van Wassenhove, L., Gabel, H.L., Weaver, P.M., 1996. Environmental life cycle optimization model for the European pulp and paper industry. Omega 24 (6), 615–629.
Bojarski, A.D., Laínez, J.M., Espuña, A., Puigjaner, L., 2009. Incorporating environmental impacts and regulations in a holistic supply chains modeling: an LCA approach. Comput. Chem. Eng. 33 (10), 1747–1759.
Cardona-Valdés, Y., Álvarez, A., Ozdemir, D., 2011. A bi-objective supply chain design problem with uncertainty. Transport. Res. Part C: Emerg. Technol. 19 (5), 821–832.
Cheshmehgaz, H.R., Desa, M.I., Wibowo, A., 2013. A flexible three-level logistic network design considering cost and time criteria with a multi-objective evolutionary algorithm. J. Intell. Manuf. 24 (2), 277–293.
Chibeles-Martins, N., Pinto-Varela, T., Barbósa-Póvoa, A.P., Novais, A.Q., 2012. A simulated annealing algorithm for the design and planning of supply chains with economic and environmental objectives. Comput. Aided Chem. Eng. 30, 21–25.
Colombo, G., Mumford, C., 2005. Comparing algorithms, representations and operators for the multi-objective knapsack problem. The 2005 IEEE Congress on Evolutionary Computation, vol. 2. IEEE, pp. 1268–1275.
Corne, D., Knowles, J., Oates, M., 2000. The Pareto envelope-based selection algorithm for multiobjective optimization. In: Schoenauer, M. et al. (Eds.), Parallel Problem Solving from Nature PPSN VI, Ser. Lecture Notes in Computer Science, vol. 1971. Springer, Berlin/Heidelberg, pp. 839–848. http://dx.doi.org/10.1007/3-540-45356-3_82.
Daskin, M.S., 1995. Network and Discrete Location: Models, Algorithms and Applications. Wiley-Interscience Publication, New York.
Deb, K., 2001. Multi-Objective Optimization using Evolutionary Algorithms. John Wiley & Sons.
Deb, K., Agrawal, S., Pratap, A., Meyarivan, T., 2000. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimisation: Nsga-II. In: Proceedings of the 6th International Conference on Parallel Problem Solving from Nature, Ser. PPSN VI. Springer-Verlag, London, UK, pp. 849–858, <http://portal.acm.org/citation.cfm?id=645825.668937>.
DEFRA, 2005. Department for Environment, Food & Rural Affairs (DEFRA). Guidelines for Company Reporting on Greenhouse Gas Emissions. London.
Delmaire, H., Diaz, J.A., Fernandez, E., 1999. Reactive GRASP and tabu search based heuristics for the single source capacitated plant location problem. INFOR, Can. J. Oper. Res. Inform. Process. 37 (3), 194–225.
Ding, H., Benyoucef, L., Xie, X., 2006. A simulation-based multi-objective genetic algorithm approach for networked enterprises optimization. Eng. Appl. Artif. Intell. 19 (6), 609–623.
Drezner, Z., Hamacher, H.W., 2002. Facility Location: Applications and Theory. Springer, Berlin.
Edwards, J., B., McKinnon, A.C., Cullinane, S.L., 2010. Comparative analysis of the carbon footprints of conventional and online retailing: a "last mile" perspective. Int. J. Phys. Distrib. Logist. Manage. 40 (1–2), 103–123.
Farahani, R.Z., SteadieSei, M., Asgari, N., 2010. Multiple criteria facility location problems: a survey. Appl. Math. Model. 34 (7), 1689–1709.
Fisher, M., 1981. The Lagrangian relaxation method for solving integer programming problems. Manage. Sci. 27 (1), 1–18.
Fisher, M., 1985. An applications oriented guide to Lagrangian relaxation. Interfaces 15 (2), 10–21.
Florios, K., Mavrotas, G., Diakoulaki, D., 2010. Solving multiobjective, multiconstraint knapsack problems using mathematical programming and evolutionary algorithms. Eur. J. Oper. Res. 203 (1), 14–21.
Gen, M., Cheng, R., 1997. Genetic Algorithms and Engineering Design. John Wiley & Sons, New York.
Ghiani, G., Laporte, G., Musmanno, R., 2004. Designing the logistics network. Introduction to Logistics Systems Planning and Control, Ser. Wiley Interscience Series in Systems and Optimization. John Wiley & Sons, pp. 88–120 (Chapter 3).
Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA.
Harris, I., Mumford, C., Naim, M., 2009. The multi-objective uncapacitated facility location problem for green logistics. In: IEEE Congress on Evolutionary Computation, 2009 (CEC '09), Trondheim, Norway, 18–21 May 2009. IEEE, Los Alamitos, CA, pp. 2732–2739.
Harris, I., Naim, M., Palmer, A., Potter, A., Mumford, C., 2011a. Assessing the impact of cost optimization based on infrastructure modelling on CO_2 emissions. Int. J. Prod. Econ. 131 (1), 313–321.
Harris, I., Mumford, C., Naim, M., 2011b. An evolutionary bi-objective approach to the capacitated facility location problem with cost and CO_2 emissions. In: Genetic and Evolutionary Computation, GECCO 2011. ACM.
Harrisdata, 2011. <http://users.cs.cf.ac.uk/C.L.Mumford/Research%20Topics/FLP/papers/data/>.
Harrison, G.P., Maclean, E.J., Karamanlis, S., Ochoa, L.F., 2010. Life cycle assessment of the transmission network in Great Britain. Energy Policy 38 (7), 3622–3631.
Hiremath, N.C., Sahu, S., Tiwari, M.K., 2013. Multi objective outbound logistics network design for a manufacturing supply chain. J. Intell. Manuf. 24 (6), 1071–1084.
Holmberg, K., Ronnqvist, M., Yuan, D., 1999. An exact algorithm for the capacitated facility location problems with single sourcing. Eur. J. Oper. Res. 113 (3), 544–559.
Hugo, A., Pistikopoulos, E.N., 2005. Environmentally conscious long-range planning and design of supply chain networks. J. Cleaner Prod. 13, 1471–1491.
IBM, 2009. ILOG® CPLEX® v12.1. User's Manual for CPLEX.
Javanshir, H., Ebrahimnejad, S., Nouri, S., 2012. Bi-objective supply chain problem using MOPSO and NSGA-II. Int. J. Ind. Eng. Comput. 3 (4), 681–694.
Jornsten, K., Nasberg, M., 1986. A new Lagrangian relaxation approach to the generalized assignment problem. Eur. J. Oper. Res. 27 (3), 313–323.

Khoo, H.H., Bainbridge, I., Spedding, T.A., Taplin, D.M., 2001. Creating a green supply chain. Greener Manage. Int. 35, 71–88.
Klose, A., Drexl, A., 2005. Facility location models for distribution system design. Eur. J. Oper. Res. 162 (1), 4–29.
Konak, A., Coit, D.W., Smith, A.E., 2006. Multi-objective optimization using genetic algorithms: a tutorial. Reliab. Eng. Syst. Saf. 91 (9), 992–1007.
Latha Shankar, B., Basavarajappa, S., Chen, J.C.H., Kadadevaramath, R.S., 2013. Location and allocation decisions for multi-echelon supply chain network – a multi-objective evolutionary approach. Expert Syst. Appl. 40 (2), 551–562.
Lee, D., Dong, M., 2008. A heuristic approach to logistics network design for end-of-lease computer products recovery. Transport. Res. Part E: Logist. Transport. Rev. 44 (3), 455–474.
Liao, S., Hsieh, C., Lin, Y., 2011a. A multi-objective evolutionary optimization approach for an integrated location-inventory distribution network problem under vendor-managed inventory systems. Ann. Oper. Res. 186 (1), 213–229.
Liao, S.H., Hsieh, C.L., Lai, P.J., 2011b. An evolutionary approach for multi-objective optimization of the integrated location-inventory distribution network problem in vendor-managed inventory. Expert Syst. Appl. 38 (6), 6768–6776.
Meade, L., Sarkis, J., 1998. Strategic analysis of logistics and supply chain management systems using the analytical network process. Transport. Res. Part E: Logist. Transport. Rev. 34 (3), 201–215.
Michalewicz, Z., 1996. Genetic Algorithms + Data Structures = Evolution Programs, third ed. Springer-Verlag, New York.
Min, H., Melachrinoudis, E., 1999. The relocation of a hybrid manufacturing/distribution facility from supply chain perspectives: a case study. Omega 27 (1), 75–85.
Mirchandani, P., Francis, R., 1990. Discrete Location Theory. John Wiley & Sons.
Mumford, C.L., 2004. Simple population replacement strategies for a steady-state multi-objective evolutionary algorithm. In: Deb, K. et al. (Eds.), Genetic and Evolutionary Computation, GECCO 20 04, Ser. Lecture Notes in Computer Science, vol. 3102. Springer-Verlag, pp. 1389–1400.
Nozick, L., Turnquist, M., 1998. Integrating inventory impacts into a fixed-charge model for locating distribution centers. Transport. Res. Part E: Logist. Transport. Rev. 34 (3), 173–186.
OR-Library, Beasley, J.E., 2009. <http://people.brunel.ac.uk/~mastjjb/jeb/info.html> > (accessed 11.03.09).
Owen, S.H., Daskin, M.S., 1998. Strategic facility location: a review. Eur. J. Oper. Res. 111, 423–447.
Paksoy, T., Pehlivan, N.Y., Özceylan, E., 2012. Application of fuzzy optimization to a supply chain network design: a case study of an edible vegetable oils manufacturer. Appl. Math. Model. 36 (6), 2762–2776.
Pati, R.K., Vrat, P., Kumar, P., 2008. A goal programming model for paper recycling system. Omega 36, 405–417.
Pishvaee, M.S., Razmi, J., 2012. Environmental supply chain network design using multi-objective fuzzy mathematical programming. Appl. Math. Model. 36 (8), 3433–3446.
Quariguasi Frota Neto, J., Bloemhof-Ruwaard, J.M., van Nunen, J.A.E.E., van Heck, E., 2008. Designing and evaluating sustainable logistics networks. Int. J. Prod. Econ. 111 (2), 195–208.
Sheu, J., Chou, Y., Hu, C., 2005. An integrated logistics operational model for green-supply chain management. Transport. Res. Part E: Logist. Transport. Rev. 41 (4), 287–313.
Snyder, L., 2006. Facility location under uncertainty: a review. IIE Trans. 38 (7), 547–564.
Ubeda, S., Arcelus, F.J., Faulin, J., 2011. Green logistics at Eroski: a case study. Int. J. Prod. Econ. 131 (1), 44–51.
Valenzuela, C., 2002. A simple evolutionary algorithm for multi-objective optimization (SEAMO). In: Proceedings of the 2002 Congress on Evolutionary Computation. IEEE, pp. 717–722.
Villegas, J.G., Palacios, F., Medaglia, A.L., 2006. Solution methods for the bi-objective (cost-coverage) unconstrained facility location problem with an illustrative example. Ann. Oper. Res. 147 (1), 109–141.
Wang, F., Lai, X., Shi, N., 2011. A multi-objective optimization for green supply chain network design. Decis. Support Syst. 51 (2), 262–269.
Xu, J., Liu, Q., Wang, R., 2008. A class of multi-objective supply chain networks optimal model under random fuzzy environment and its application to the industry of Chinese liquor. Inform. Sci. 178 (8), 2022–2043.
You, F., Tao, L., Graziano, D.J., Snyder, S.W., 2012. Optimal design of sustainable cellulosic biofuel supply chains: multiobjective optimization coupled with life cycle assessment and input–output analysis. AIChE J. 58 (4), 1157–1180.
Zitzler, E., 1999. Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications. Ph.D. Thesis. Shaker Verlag, Aachen, Germany, December.
Zitzler, E., Laumanns, M., Thiele, L., 2001. Spea2: Improving the Strength Pareto Evolutionary Algorithm. Swiss Federal Institute of Technology (ETH), Tech. Rep. 103.