OmniaScience

Journal of Industrial Engineering and Management
JIEM, 2022 – 15(4): 552-565 – Online ISSN: 2013-0953 – Print ISSN: 2013-8423
https://doi.org/10.3926/jiem.3813

A Simulation-Based Optimisation for the Stochastic Green Capacitated *p*-Median Problem

Arif Imran[1], Eko Wahyu Utomo[2], Fadillah Ramadhan[1], Arie Desrianty[1], Yanti Helianty[1], Fifi Herni Mustofa[1]

[1]*Department of Industrial Engineering, Institut Teknologi Nasional Bandung (Indonesia)*
[2]*Department of Civil Engineering, Universitas Langlangbuana, PDTS, Universitas Sultan Agung (Indonesia)*

imran@itenas.ac.id, utomoeko042@gmail.com, f.ramadhan@itenas.ac.id, adesrianty@itenas.ac.id, yanti@itenas.ac.id, fifi@itenas.ac.id

Received: October 2021
Accepted: July 2022

Abstract:

Purpose: This paper aims to propose a new model called the stochastic green capacitated *p*-median problem with a simulation-based optimisation approach. An integer linear programming mathematical model is built considering the total emission produced by vehicles and the uncertain parameters including the travel cost for a vehicle to travel from a particular facility to a customer and the amount of CO_2 emissions produced. We also develop a simulation-based optimisation algorithm for solving the problem.

Design/methodology/approach: The authors proposed new algorithms to solve the problem. The proposed algorithm is a hybridisation of Monte Carlo simulation and a Variable Neighbourhood Search matheuristic. The proposed model and method are evaluated using instances that are available in the literature.

Findings: Based on the results produced by the computational experiments, the developed approach can obtain interesting results. The obtained results display that the proposed method can solve the problems within a short computational time and the solutions produced have good quality (small deviations).

Originality/value: To the best of our knowledge, there is no paper in the previous literature investigating the simulation-based optimisation for the stochastic green capacitated *p*-median problem. There are two main contributions in this paper. First, to build a new model for the capacitated *p*-median problem taking into account the environmental impact. Second, to design a simulation-based optimisation approach to solve the stochastic green capacitated *p*-median problem incorporating VNS-based matheuristic and Monte Carlo simulation.

Keywords: stochastic, the capacitated *p*-median problem, green logistic, VNS

To cite this article:

Imran, A., Utomo, E.W., Ramadhan, F., Desrianty, A., Helianty, Y., & Mustofa, F.H. (2022). A simulation-based optimisation for the stochastic green capacitated *p*-median problem. *Journal of Industrial Engineering and Management*, 15(4), 552-565. https://doi.org/10.3926/jiem.3813

1. Introduction

The *p*-median problem (PMP) formulation was developed by ReVelle and Swain (1970), where the problem is to select the optimal location of n facilities from m potential sites. The objective of the problem is to minimise the total distances between customers and facilities, which is known as the minisum location problem. The PMP is assumed that the facilities have unlimited capacity meaning that the customers will be assigned to the nearest facilities. This problem is considered as NP-hard, which is hard to solve (Kariv & Hakimi, 1979). In contrast to the PMP, the CPMP (capacitated *p*-median problem) considers a fixed capacity for each potential facility. An open facility must satisfy customer demand without exceeding its capacity. The CPMP has a higher degree of complexity than the PMP due to this capacity constraint. Therefore, the CPMP is considered as NP-hard (Garey & Johnson, 1990).

Currently, the concern on the environmental impact of business operations is increasing. Many companies now realise that carbon emissions produced by their operations need to be reduced as the emissions have a significant impact on global warming and severe negative effects on business and society. In this study, to measure the environmental impact, we consider the amount of CO_2 emissions produced when transporting products from the facilities to customers. According to The World Meteorological Organization (2009), CO_2 is the single most human-emitted greenhouse gas emission accounting for about 63.5% of the total global warming. Moreover, CO_2 is a very popular environment index and also easily measured (Wang, Lai & Shi, 2011). Another decision that has to be made beside the location of the facility is to determine the type of vehicle used by each facility considering CO_2 emissions produced. Therefore, here, we develop a problem called the green capacitated *p*-median problem (GCPMP).

In many real case applications, some uncertain parameters need to be considered to get the best location for the facilities. Those parameters include the travel cost from a facility to a customer and the amount of CO_2 emissions produced. To deal with the problem with uncertainty, a simulation-based optimisation is proposed where hybridisation of Monte Carlo simulation and a Variable Neighbourhood Search-based algorithm is developed. To the best of our knowledge, there is no paper in the previous literature investigating the simulation-based optimisation for the stochastic GCPMP.

There are two main contributions in this paper. First, to build a new model for the capacitated *p*-median problem taking into account the environmental impact. Second, to design a simulation-based optimisation approach to solve the stochastic green capacitated *p*-median problem incorporating VNS-based matheuristic and Monte Carlo simulation.

This paper is organised as follows. In Section 2, a review of the previous studies related to the CPMP is presented. Section 3 describes the new CPMP mathematical models along with the new deterministic GCPMP. In Section 4, the ingredient of the proposed method for the stochastic GCPMP is given. Section 5 provides computational results using dataset available in the literature. The last section summarises our findings and presents some research avenues for future research.

2. Literature Review

The *p*-median problem is one of the models that involves the location of n facilities on the networks and it is one of the popular location problems in the literature. This classical problem is commonly identified to minimize the total distance in serving all demands. Mulvey and Beck (1984) can be considered as the earliest ones who work on the CPMP. They proposed two algorithms to solve the capacitated clustering problems. Osman and Christofides (1994) put forward a hybridisation of simulated annealing and tabu search method to solve the CPMP. A bionomic algorithm was designed by Maniezzo, Mingozzi and Baldacci (1998) as a local search to address the CPMP. A set partitioning formulation method for the CPMP was developed by Baldacci, Hadjiconstantinou, Maniezzo and Mingozzi (2002). Lorena and Senne (2004) developed a column-generation technique to deal with the CPMP. A hybridisation of the Adaptive Memory Programming and the Greedy Random Adaptive Search Procedure for tackling the CPMP was proposed by Ahmadi and Osman (2005).

Scheuerer and Wendolsky (2006) solved the CPMP by using the scatter search heuristic. A hybridisation of scatter search algorithm and path relinking was developed by Díaz and Fernàndez (2006), whereas Fleszar and Hindi (2008) built an effective VNS for tackling the CPMP. A hybrid heuristic which is referred to it as clustering search was introduced by Chaves, Correa and Lorena (2007). Boccia, Sforza, Sterle and Vasilyev (2008) put forward an efficient cutting plane algorithm to reduce the integrality gap in solving the CPMP. Genetic algorithms and harmony search method were proposed by Landa-Torres, Ser, Salcedo-Sanz, Gil-Lopez, Portilla-Figueras and Alonso-Garrido (2012) to address the CPMP.

A heuristic approach based on the local branching and relaxation induced neighbourhood search methods was developed by Ghoseiri and Ghannadpour (2013). Yaghini, Karimi and Rahbar (2013) put forward an algorithm that hybridises a tabu search metaheuristic and a cutting-plane neighbourhood structure for the CPMP. A three-stage matheuristic method introduced by Stefanello, Araujo and Muller (2014) to solve the CPMP. Three methods were applied by El Amrani, Benadaa and Gendron (2016), namely a branch and cut algorithm, large neighbourhood search, and greatest customer demand first to address the CPMP considering the budget constraint into the problem. Irawan, Imran and Luis (2017) proposed the bi-objective CPMP with the fixed cost for opening the new facility and its capacities that can be utilized by potential facilities. The problem is solved by using a compromise programming approach. CPMP has many various types and different characteristics for each case. There are some studies that identify the stochastic aspect in the CPMP model, and others try to find a best solution of dynamic and complex CPMP.

The stochastic CPMP considering the environmental aspect which has not been covered by the papers cited above is addressed in this study. We also propose a simulation-based optimisation to deal with such a problem. Here, the hybridisation of Monte Carlo simulation and VNS-based matheuristic are proposed.

3. The Green Capacitated *p*-Median Problem (GCPMP)

The objective of CPMP is to locate p facilities (medians) so as to minimize the sum of the distances from each demand point to its nearest facility. Bramel and Simchi-Levi (1995), Klein and Aronson (1991) and Mulvey and Beck (1984) are among the first who investigate this problem. The proposed mathematical model for the deterministic green capacitated *p*-median problem (GCPMP) is given in this section. The model is developed from the mathematical model of the classical CPMP.

In the new mathematical model of the GCPMP, the presence of truck/vehicle types is considered where set V is used to represent the set of vehicles. Here, the vehicle type used by each open facility $j \in J$ the decision variable (Y_{jv}). Note that an open facility only uses one type of vehicle. Each vehicle $v \in V$ has a different travel cost per unit distance (C_v) and produces a different amount of CO_2 emissions (e_v). An additional constraint is added to ensure that total emissions produced do not exceed e^{max}. In addition to opening facility decision ($\hat{Y}_j$), the model aims to solve the allocation problem to determine whether open facility $j \in J$ serves customer $i \in I$ using vehicle $v \in V$ (X_{vij}). In summary, the following notations for sets, decision variables, and parameters are used:

Set

I customer set ($i \in I = \{1, \ldots, n\}$, $n = |I|$)

J potential site set ($j \in J = \{1, \ldots, m\}$, $m = |J|$)

V vehicle (truck) types

Parameters

d_{ij} the distance between customer $i \in I$ and facility $j \in J$

w_i the demand from customer $i \in I$

b_j the facility capacity that is located at the site $j \in J$

p the number of open facilities

c_v the travel cost per unit distance to deliver one unit product using vehicle $v \in V$

e_v the amount of CO_2 emissions produced per unit distance caused by delivering one unit product using vehicle $v \in V$

e^{max} the maximum CO_2 emissions produced

Decision Variables

X_{vij} = 1, if facility j serves customer i using vehicle v;
0, otherwise

Y_{jv} = 1, if the facility is situated at site j using vehicle v;
0, otherwise

$\hat{Y}_j$ = 1, if the facility is situated at site j
0, otherwise

The proposed model aims to find the optimal solution for both the vehicle and facilities' location. The problem can be modelled as follows:

Minimise

$$\sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\left(w_i \cdot d_{ij} \cdot c_v \cdot X_{vij}\right) \tag{1}$$

Subject to

$$\sum_{j\in J}\sum_{v\in V} X_{vij} = 1, \quad \forall i \in I \tag{2}$$

$$\sum_{v\in V} Y_{jv} \leq \hat{Y}_j, \quad \forall j \in J \tag{3}$$

$$\sum_{j\in J} \hat{Y}_j = p \tag{4}$$

$$\sum_{i\in I}\sum_{v\in V} X_{vij} \cdot w_i \leq b_j \cdot \hat{Y}_j, \quad \forall j \in J \tag{5}$$

$$X_{vij} - Y_{jv} \leq 0, \quad \forall i \in I, j \in J, v \in V \tag{6}$$

$$\sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\left(w_i \cdot e_v \cdot d_{ij} \cdot X_{vij}\right) \leq e^{max} \tag{7}$$

$$Y_{jv} \in \{0,1\}, \forall j \in J, v \in V \tag{8}$$

$$X_{vij} \in \{0,1\}, \quad \forall i \in I, j \in J, v \in V \tag{9}$$

$$\hat{Y}_j \in \{0,1\}, \quad \forall j \in J \tag{10}$$

The objective function (1) describes the total travel cost from demand points to their corresponding facilities. Constraints (2) make sure each customer demand must be fulfilled by only one facility. Constraints (3) ensure that facility only uses one vehicle type. Constraint (4) guarantees that p open facilities are opened. Constraints (5) ensure that the capacity of an open facility will not be exceeded. Constraints (6) ensure that each customer can only be assigned to an open facility. Constraint (7) makes sure that the total CO_2 emissions produced do not exceed e^{max}. The last three constraints state the decision variables binary conditions.

4. Simulation-Based Optimisation for the Stochastic GCPMP

A simulation-based optimisation approach is developed to deal with the stochastic GCPMP. In this study, two parameters are considered as stochastic parameters which are as follows:

- c_v the travel cost per unit distance to deliver one unit product using vehicle $v \in V$
- e_v the amount of CO_2 emissions produced per unit distance caused by delivering one unit product using vehicle $v \in V$

Those parameters are assumed to follow a normal distribution and the realisations of that assumption have a good outcome. In the routing problem, normal distribution is widely used to estimate the travel time or travel cost (Li, Tian & Leung, 2010; Shen, Xu, Wu & Ni, 2019; Bakach, Campbell, Ehmke & Urban, 2021). In our approach, the VNS-based matheuristic technique given in the next subsection is applied to address the deterministic problem to determine the facilities' location and customers' allocation, together with their responding vehicle. Here, the parameter that has a stochastic characteristic is transformed into deterministic value when solving this deterministic problem. Once the facility configuration has been obtained, the Monte Carlo simulation is implemented to attain the total cost estimation, including the stochastic parameters.

The main procedure of the simulation-based optimisation method is presented in Figure 1, where hybridisation of a VNS-based matheuristic and Monte Carlo simulation is put forward. In the initial step, parameters θ, B, S and L are defined. Parameter θ represents the θ%-quantile total cost data in the simulation that will be used to determine the expected total cost. A higher value of θ will generate a robust solution for the worse scenario.

Initialisation

Define θ, B, S and L. Set $Z^* \leftarrow \infty$

Phase 1

For $\alpha = 1$ to B do the following steps:

1. If $\alpha = 1$ then
   Use the mean value to estimate the travel cost (c_v) and the emissions produced (e_v)
   Else
   - Set $\lambda = (\alpha - 1) \cdot 10$
   - Use the λ%-quantile data that is generated from Monte Carlo simulation in the iteration $[\alpha - 1]$ to estimate travel cost (c_v) and emissions produced (e_v) parameters

   End if
2. Solve the deterministic GCPMP using the VNS-based matheuristic method given in Figure 2 with deterministic parameters. The facility configuration (Y_{jv}) and customers allocation (X_{vij}) are then obtained.
3. Run Monte Carlo S times using the facility configuration (Y_{jv}) and customers allocation (X_{vij}) by calling procedure MC_Sim (S, θ and Z) given in Figure 4. Record the total cost (Z) based on the θ%-quantile total cost data.
4. If $Z < Z^*$ then set $Z^* = Z$, $Y^*_{jv} \leftarrow Y_{jv}$ and $X^*_{vij} \leftarrow X_{vij}$

Phase 2

Run Monte Carlo L times using Y^*_{jv} and X^*_{vij} obtained from the previous stage (i.e. MC_Sim (L, θ and Z)). Record Z^* (the final total cost value) and take Y^*_{jv} and X^*_{vij} as the facility configuration and the customers' allocation in the final robust solution.

Figure 1. The proposed simulation-based optimisation for the stochastic GCPMP

The algorithm of Figure 1 comprises two phases which are developed based on the literature of Irawan, Eskandarpour, Ouelhadj and Jones (2021). The initial stage is an iteration process where the proposed VNS-based matheuristic method is used to find the facility configuration (Y_{jv}). Here, the stochastic parameters c_v and e_v are treated as deterministic parameters. The first iteration of Phase 1 determines those parameters based on their average values. In the remaining iterations, the values of those parameters are calculated based on the λ%-quantile data produced from Monte Carlo simulation in the previous iteration. The value of λ is adjusted systematically using the following formula: $\lambda = (\alpha - 1)10$. Once the facility configuration (Y_{jv}) has been fixed, the Monte Carlo simulation is conducted to determine the total cost (Z), including the stochastic parameters. The Monte Carlo simulation procedure is given in the next subsection. In Phase 1, The simulation is executed S times which can be seen as short simulation (e.g., S = 10,000). This short simulation procedure is repeated B times, and the best facility configuration (Y_{jv}) that provides the lowest total cost (Z^*) is chosen. As the simulation-based optimisation approach needs to solve the deterministic model iteratively (several times), therefore, a VNS-based matheuristic method is developed to solve the problem within a short time while delivering good quality solution.

In Phase 2, Monte Carlo simulation is executed based on (Y_{jv}) to determine the final total cost (Z^*). Here, a long simulation is performed to obtain the expected total cost. The number of iterations (L) for this long simulation is set to a high value (e.g., L = 100,000).

4.1. The VNS Based Matheuristic

Brimberg and Mladenović (1995) presented VNS to tackle location-allocation problems that have a continuous characteristic. However, Variable Neighborhoud Search (VNS) was first formally formulated by Hansen and Mladenović (1997) when they apply it to address the *p*-median problem. More information on VNS types and applications can be found in Hansen, Mladenovic and Perez (2008), and Hansen and Mladenović (2001). Basically, VNS consists of two elements, namely neighbourhood search and local search. The neighbourhood search objective is to help the search process escape from the local optima while the local search tries to obtain local optimality. If the search process cannot find any improvement, the neighbourhood search applies the next neighbourhood (a larger one), otherwise, it reverts to the smaller neighbourhood. In this study, VNS-based matheuristic is constructed based on the one proposed by Irawan et al. (2017). The VNS matheuristic main steps can be seen in Figure 2.

The main objective of the proposed method is to obtain the best solution (S) that yields the smallest total cost (z) together with the vehicle configuration (K). The parameters needed by the proposed matheuristic approached are defined in the Initialisation step. The parameters include T (the number of iterations) to solve the reduced problems, µ as the number of reduced potential facilities, the %Gap for CPLEX for solving the reduced problems (τ) and k_{max} as the VNS last neighbourhood.

The next step is an iteration approach and aims to produce a good initial solution. A reduced problem set is constructed and solved by the exact method (CPLEX). The number of reduced problems is the same as the number of iterations (T) that has been set up in the initialisation step. In a reduced problem, the number of potential facility sites is reduced to (μ) locations out of m locations. Here, μ facility sites consist of randomly chosen facilities and set S, which is the promising sites obtained from the previous iterations. The reduced problem consists of n customers, and (μ) potential facility sites are then solved by CPLEX and will be terminated once it reached τ% gap. S and K denote the facility sites configuration and their corresponding vehicle. The solution obtained is then used to successive iteration as the promising potential sites part. Repeat the process T times, and in the VNS algorithm, the best solution produced by this step will be used. This solution method has good performance for solving large *p*-median and *p-centre* problem (Irawan, Salhi & Scaparra, 2014; Irawan & Salhi, 2013; Irawan, Salhi & Drezner, 2016).

Step 5 is first performed by calculating the distance criterion $\hat{d}_j, j \in S'$. Parameter $\hat{d}_j$ is determined based on the average distance between facility j and the customers allocated to this facility. A facility, say facility $\hat{j}$ $k \leq k_{max}$ is randomly chosen, where facility j is assigned to a customer located in site $\hat{j}$. If the distance between facility j and $\hat{j}$ ($d_{j\hat{j}}$) is less than $\hat{d}$ then facility j is replaced by facility $\hat{j}$. This is done to avoid the solution perturbed too much. This procedure is conducted k times. Then, the allocation problem represented by Equations (11-18) is solved using CPLEX based on solution S'. The objective function value (z') and vehicle configuration (set K') are obtained.

VNS-based matheuristics procedure

Initialisation step

1. Define T, μ, τ and (k_{max}). Set $z = \infty$ and $S \leftarrow \emptyset$.

Iterative step

2. Do step a to c T times:
   a. Let $\check{J}$ $(|\check{J}| = \mu)$ a subset of potential facility sites set J $(\check{J} \subseteq J)$ where $(\mu - |S|)$ sites are randomly chosen from m potential facility sites. Include the promising facility locations (S) in set $\check{J}$.
   b. Solve the reduced problems that consist of n customers and (μ) potential sites using the CPLEX until τ % gap is reached. Let z' be its objective function value with S' and K' as sets of the obtained facility location configuration and their corresponding vehicle respectively.
   c. If $z' = z$ then set $z = z'$, $S \leftarrow S'$ and $K \leftarrow K'$.
3. Update $z' = z$, $S' \leftarrow S$, and $K' \leftarrow K$

VNS step

4. Set $k = 1$
5. Shaking Procedure.
   a. Calculate the distance criterion $\hat{d}_j, j \in S'$
   b. Do step i to iii k times:
      i. Pick a potential facility randomly, say facility $\hat{j}(\hat{j} \notin S')$, where a customer located in site $\hat{j}$ is allocated to facility j in Solution S'
      ii. If $d_{j,\hat{j}} > \hat{d}_j$ then go back to Step 5b(i)
      iii. Replace facility j with facility j in the solution S'
   c. Solve the allocation problem (11) – (18) using CPLEX based on S' where z' and set K' will be obtained.
6. Local search
   Run local search which is described in Figure 3 with z', S' and K' as inputs and outputs.
7. Move or Not
   If $z' < z$ then
      Update $k = 1$, $z = z'$, $S \leftarrow S'$, and $K \leftarrow K'$
   Else
      Update $k = k + 1$, $z' = z$, $S' \leftarrow S$, and $K' \leftarrow K$
8. If $k \leq k_{max}$ revert to Step 5
   Else Stop and Return z, S, and K.

Figure 2. The VNS-based matheuristic procedure

Minimise

$$\sum_{i \in I} \sum_{j \in S'} \sum_{v \in V} (w_i \cdot d_{ij} \cdot c_v \cdot X_{vij}) \tag{11}$$

Subject to

$$\sum_{j \in S'} \sum_{v \in V} X_{vij} = 1, \quad \forall i \in I \tag{12}$$

$$\sum_{v \in V} Y_{jv} = 1, \quad \forall j \in S' \tag{13}$$

$$\sum_{i \in I} \sum_{v \in V} X_{vij} \cdot w_i \leq b_j, \quad \forall j \in S' \tag{14}$$

$$X_{vij} - Y_{jv} \leq 0, \quad \forall i \in I, j \in S', v \in V \tag{15}$$

$$\sum_{i \in I} \sum_{j \in S'} \sum_{v \in V} \left(w_i \cdot e_v \cdot d_{ij} \cdot X_{vij} \right) \leq e^{max} \tag{16}$$

$$Y_{jv} \in \{0,1\}, \quad \forall j \in S', v \in V \tag{17}$$

$$X_{vij} \in \{0,1\}, \quad \forall i \in I, j \in S', v \in V \tag{18}$$

In Step 6, we propose a local search to improve the solution quality, which is described in the next subsection. Step 7 is conducted where if the local search procedure can improve the solution, the smallest neighbourhood (k=1) will be applied; otherwise, a larger neighbourhood is systematically used ($k = k + 1$). Once the value of k reaches k_{max}, the algorithm terminates and the best results are taken.

4.2. The Local Search

The procedure of the local search used in the VNS algorithm is explained in this section. Figure 3 shows the procedure of the proposed local search where the best improvement strategy is used. The main objective of the proposed algorithm is to seek a location site (in set J) to be swapped with a facility site used in the current solution (set S). Once the best improvement is found, the swap process is performed, and the new vehicle configuration is obtained. The distance criterion between facilities ($\hat{d}_j$) in the current solution (S) is first determined as a criterion to see if we can swap a facility site in the current solution with other potential facility site or not. Here, we do not want to swap a facility in the current solution with a potential facility which is very far from that facility.

Procedure Local Search (z, S, and K)

1. Calculate the distance criterion $\hat{d}_j, j \in S$
2. Set Improve = false and Sav = 0
3. For each potential facility $\hat{j} \in J, \hat{j} \notin S$, do:
   For each facility $j \in S$ (current solution) do the following procedure:
   a. If in current solution a customer located in site $\hat{j}$ is not allocated to facility j then continue (skip following steps under loop j)
   b. If $d_{j,\hat{j}} > \hat{d}_j$ then continue (skip following steps under loop j)
   c. In solution S, replace facility j with $\hat{j}$
   d. Solve the allocation problem (19) – (22) using CPLEX based on solution S. Let z' be the objective function value
   e. If $z - z' >$ Sav then
      - Update Sav $= z - z'$ together with $j_{out} = j$ and $j_{in} = \hat{j}$
      - Update Improve = true
   f. In solution S, replace facility $\hat{j}$ with j
   End for j
   End for $\hat{j}$
4. If (Improve = true) do the followings:
   - In solution S, replace facility j_{out} with j_{in}
   - Solve the allocation problem (11) – (18) using CPLEX based on solution S.
   - Update objective function value z, set K and the distance criterion $\hat{d}_j, j \in S$
   - Go to Step 2
   Else Stop and Return z, S, and K.

Figure 3. The local search main steps

We also evaluate each potential facility to be inserted in the solution by solving the allocation problem presented in Equations (19-22) using CPLEX (Step 3d). Here, when facility j is replaced by facility $\hat{j}$, the vehicle type used by facility $\hat{j}$ is based on the one used by facility j. In other words, there is no change in vehicle configuration.

Minimise

$$\sum_{i\in I}\sum_{j\in S}\left(w_i \cdot d_{ij} \cdot c^v \cdot X_{ij}^v\right) \tag{19}$$

Subject to

$$\sum_{j\in S} X_{ij}^v = 1, \quad \forall i \in I \tag{20}$$

$$\sum_{i\in I} X_{ij}^v \cdot w_i \le b_j, \quad \forall j \in S \tag{21}$$

$$X_{ij}^v \in \{0,1\}, \quad \forall i \in I, j \in S \tag{22}$$

Once the best improvement is found (Step 4), the allocation problem (11-18) is solved using CPLEX. We may have a different vehicle configuration, and the maximum emissions constraints are met. We repeat the procedure until there is no improvement found.

4.3. Monte Carlo Simulation

In this subsection, the Monte Carlo simulation approach is described. Monte Carlo simulation is a repeated process where, for each replication, the random number is generated to represent the uncertainty or stochastic parameters (travel cost and emission produced). For each replication, as the emission produced per vehicle per unit distance is randomly generated, there is a chance that the solution of the problem is not feasible because the total emissions produced is greater than the maximum emissions produced allowed. Therefore, we propose an emission cost per unit for not meeting the maximum emissions, which is considered as the recourse cost. Here, we add a parameter represented by c_e. It represents the emission cost of one unit of emissions produced higher than the maximum one.

Once the facility location, customer allocation and the vehicle configurations are fixed (which is represented by set S and K respectively), the decision variables Y_{jv} and X_{vij} are treated as a parameter. Here, we transform variables Y_{jv} and X_{vij} to parameters $\tilde{Y}_{jv}$ and $\tilde{X}_{vij}$, respectively. The total cost for random values of stochastic parameters can be obtained by calculating the total travel cost and the emission cost if it occurs. The total cost, including the simulation emission cost, is formulated as follows:

$$\begin{aligned} Z = \sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\left(w_i \cdot d_{ij} \cdot c_v \cdot \tilde{X}_{vij}\right) + \\ c_e \cdot \max\{(\textstyle\sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\left(w_i \cdot e_v \cdot d_{ij} \cdot \tilde{X}_{vij}\right) - e_{max}),\ 0\}) \end{aligned} \tag{23}$$

Figure 4 presents the Monte Carlo procedure used for the GCPMP where the simulation is executed $\tilde{T}$ times (replications). For each replication, parameters c_v and e_v are randomly generated based on their distribution. By calculating Equation (23), the total cost considering the emission cost is determined. The expected total cost is taken from the $\gamma\%$-quantile total assignment cost (z_t).

Procedure Monte Carlo Simulation ($\tilde{T}$, γ and z^S)

1. For $t = 1$ to $\tilde{T}$ do:
   - Generate randomly parameters c_v and e_v based on their distribution
   - Compute the total cost of the t^{th} replication (z_t) by calculating Equation (23)
2. Apply the $\gamma\%$-quantile total cost data (z_t) as the expected total cost value (Z^S)

Figure 4. The procedure of Monte Carlo simulation

5. Computational Study

In this study, the proposed approach performance is examined by conducting extensive experiments. The algorithm is programmed in C++.Net 2017. The IBM ILOG CPLEX version 12.8 Concert Library is also used to get a solution with an exact method. The computational experiments are executed on a PC with an Intel Xeon W-2133 CPU @3.60 GHz processor and 64.00 GB of RAM. The dataset used to evaluate the proposed solution method is constructed based on Irawan et al. (2021) for the customer locations and Al-e-hashem and Rekik (2014) for the travel cost and emission produced by each type of vehicle. Three datasets are considered with n = 300 up to 600 (increment value is 150). The potential facilities location are located in the customer sites, i.e. $|J| = m = n$, in this computational experiment. Each potential facility has two possible vehicle types ($|V| = 2, j \in J$). The capacity of each potential facility (b_j) and the maximum total emissions produced (e^{max}) are generated differently for each instance. The value of p (the open facilities number) varies from 3 to 6 with an increment of 1. Dataset can be obtained from the first author.

In this subsection, the Monte Carlo simulation approach is described. Monte Carlo simulation is a repeated process where, for each replication, the random number is generated to represent the uncertainty or stochastic parameters (travel cost and emission produced). For each replication, as the emission produced per vehicle per unit distance is randomly generated, there is a chance that the solution of the problem is not feasible because the total emissions produced is greater than the maximum emissions produced allowed. Therefore, we propose an emission cost per unit for not meeting the maximum emissions, which is considered as the recourse cost. Here, we add a parameter represented by c_e. It represents the emission cost of one unit of emissions produced higher than the maximum one.

5.1. The Experimental Results on the Deterministic GCPMP

The experiments on the deterministic GCPMP are first conducted to evaluate the proposed VNS matheuristic. The deterministic GCPMP is tackled by two solution approaches, namely the exact method (EM) and the proposed VNS. The proposed VNS evaluation is performed by comparing the proposed approach solutions with the exact method solutions using CPLEX. The exact method (EM) implementation is conducted by using CPLEX. As the problem is relatively hard to solve by the EM, the CPLEX computational time is limited (three hours for each problem) where lower bound (LB) and upper bound (UB) are obtained. The proposed solution method performance is assessed by the deviation (Gap) between the z value obtained by the proposed VNS and the lower bound obtained by the exact method. The Gap (%) is computed as follow:

$$Gap = 100 \times \frac{z_p - LB}{z_p} \tag{24}$$

where z_p represents the objective function value with the feasible solution found by either the proposed VNS or the exact method.

The T parameter = min(10, 2p), μ = min(50,4p), τ = 2% and k_{max} = min(p, 5) were chosen based on the preliminary experiments. The T, μ, and τ values affect the initial solution quality produced. The larger the T and μ values with the smaller τ, the greater chance to get better initial solution quality. However, it will increase the computational time.

Table 1 presents experimental results on the deterministic GCPMP using the proposed VNS and the exact method (EM). The first two columns refer to the potential facilities /customers number and the open facilities number. For the EM, the table presents UB, LB, %Gap, and CPU. Similarly, the results of the proposed VNS are also given in the table representing by the objective function value (Z), %Gap and CPU time (in seconds).

Based on Table 1, CPLEX can get optimal solutions (0% gap, recorded in **bold**) only for seven instances within 3 hours. Based on the average deviation, the proposed VNS produces a smaller average gap than the exact method. The average gap of the proposed VNS is 0.4841, and the exact method average gap is 0.5034. In general, the VNS based matheuristic can be considered as the better performer for solving the GCPMP as it produced the smallest gap. It also can find a solution much faster than the EM as the EM needs more than forty times longer to solve the problem compared to the VNS based matheuristic. Therefore, the proposed VNS can be very useful to be incorporated in the proposed simulation-based optimisation method for solving the GCPMP. This approach needs an optimiser (solution method) to be executed iteratively. It can be fulfilled by introducing a powerful optimiser that runs very fast while producing good solutions. The EM is not practical as it consumes a long computational time, especially for large problems (*n* = 600).

n	*p*	EM				VNS		
		UB	LB	%Gap	CPU	Z	%Gap	CPU
300	3	1,262,725.18	1,262,725.18	**0.0000**	804	1,262,725.18	**0.0000**	28
	4	1,011,448.84	1,011,448.84	**0.0000**	706	1,011,448.84	**0.0000**	26
	5	907,078.79	907,078.79	**0.0000**	1,257	907,078.79	**0.0000**	48
	6	888,009.75	875,640.61	1.3929	10,801	885,675.77	1.1331	94
450	3	3,026,386.79	3,026,386.79	**0.0000**	5,113	3,036,122.43	0.3207	40
	4	2,368,389.88	2,368,389.88	**0.0000**	2,367	2,368,389.88	**0.0000**	54
	5	2,131,473.77	2,131,473.77	**0.0000**	2,787	2,131,473.77	**0.0000**	153
	6	1,957,662.08	1,957,662.08	**0.0000**	2,852	1,958,724.12	0.0542	92
600	3	5,287,729.68	5,179,072.92	2.0549	10,805	5,287,220.80	2.0455	68
	4	4,252,100.68	4,239,119.95	0.3053	10,804	4,252,100.68	0.3053	93
	5	3,927,965.09	3,883,596.95	1.1295	10,809	3,914,641.78	0.7930	312
	6	3,756,667.34	3,713,140.26	1.1587	10,803	3,756,629.58	1.1577	636
				0.5034	5,826		**0.4841**	137

Table 1. Experimental results on the deterministic CPMP

5.2. The Experimental Results on the Stochastic GCPMP

Computational experiments were used for evaluating the performance of the simulation-based optimisation method. Here, the instances provided in the previous subsection are used. The standard deviation of travel cost and emission produced for each vehicle is added, whereas their average data is based on data on the deterministic problem. The standard deviation of each stochastic parameter is set to 20% of the average data. In the Monte Carlo simulation, the emission cost is also added if the total emissions produced are higher than the maximum emissions allowed.

In this experiment, parameter B, S, and L are set to 10, 10000 and 100000 respectively. We varied the value of θ to 50%, 70% and 90% in the experiments conducted. It can be said that the expected total cost is obtained by using 50%, 70%, and 90% quantiles. Table 2 shows the experimental results for the stochastic GCPMP summary when the number of open facilities is set to 3 and 4. The deterministic results in Table 2 are found by solving the deterministic GCPMP presented in the previous section. The mean values are used for the stochastic parameters. The table also reveals the results of the stochastic GCPMP when 50%, 70% and 90% quantiles are applied. In Table 2, we can also see the difference (%) between the deterministic GCPMP solution and the stochastic GCPMP solution. If the value of θ increases, it can be noted that the difference increases. A difference of 1.3286%, 11.3229% and 15.9561% yielded for 50%, 70% and 90% quantiles.

n	*p*	*Z* Deterministic	Quantile (50%)		Quantile (70%)		Quantile (90%)	
		Problem	Z	Diff. (%)	Z	Diff. (%)	Z	Diff. (%)
300	3	1,262,725.18	1,293,570.13	2.3845	1,443,528.42	12.5251	1,542,802.43	18.1538
	4	1011448.842	1,018,246.44	0.6676	1,127,217.69	10.2703	1,145,993.63	11.7404
450	3	3026386.787	3,099,779.65	2.3677	3,481,866.50	13.0815	3,720,396.92	18.6542
	4	2368389.875	2,399,914.54	1.3136	2,663,932.92	11.0942	2,759,959.47	14.1875
600	3	5287220.797	5,310,051.12	0.4299	5,838,230.68	9.4380	6,264,854.08	15.6050
	4	4252100.684	4,286,759.51	0.8085	4,806,167.41	11.5282	5,147,562.70	17.3958
				1.3286		11.3229		15.9561

Table 2. The experimental results summary on the stochastic CPMP

6. Conclusions

This paper addresses the stochastic green capacitated *p*-median problem by using hybridisation of Monte Carlo simulation and a VNS based matheuristic. A mathematical model for the capacitated *p*-median problem taking into account the environmental impact is developed. As the EM experiences problems to find the optimal solutions, a simulation-based optimisation algorithm is developed. It is a hybridisation of Monte Carlo and VNS based matheuristic. The proposed algorithm includes a Monte Carlo simulation, the VNS algorithm, and the exact method. To measure the performance of the proposed algorithm, data sets from the literature were solved and the results obtained are compared with the results of the exact method. The obtained results display that the proposed method can solve the problems within a short computational time and the solutions produced have good quality (small deviations).

For future research, the following research directions might be deserving of investigation. To obtain faster results, a pure metaheuristic method can be used to solve the GCPMP. In this paper, a deterministic demand is used, in the future investigation, the formulation can be developed by considering the stochastic demand. In terms of method, instead of using a simulation-based method, a scenario based-method can be applied to get a better solution. This model can be further developed into multi-objective optimisation. Besides that, several aspects can be included for the future model, such as logistic product priority, refuelling station for the vehicle, and also perishable product condition.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

Ahmadi, S., & Osman, I.H. (2005). Greedy random adaptive memory programming search for the capacitated clustering problem. *European Journal of Operational Research*. 162(1), 30-44. https://doi.org/10.1016/j.ejor.2003.08.066

Al-e-hashem, S.M.J.M., & Rekik, Y. (2014). Multi-product multi-period Inventory Routing Problem with a transshipment option: A green approach. *International Journal of Production Economics*, 157, 80-88. https://doi.org/10.1016/j.ijpe.2013.09.005

Bakach, I., Campbell, M., Ehmke, J.F., & Urban, T.L. (2021). Solving vehicle routing problems with stochastic and correlated travel times and makespan objectives. *EURO Journal on Transportation and Logistics*, 10, 100029. https://doi.org/10.1016/j.ejtl.2021.100029

Baldacci, R., Hadjiconstantinou, E., Maniezzo, V., & Mingozzi, A. (2002). A new method for solving capacitated location problems based on a set partitioning approach. *Computers & Operations Research*, 29(4), 365-386. https://doi.org/10.1016/S0305-0548(00)00072-1

Boccia, M., Sforza, A., Sterle, C., & Vasilyev, I. (2008). A cut and branch approach for the capacitated *p*-median problem based on fenchel cutting planes. *Journal of Mathematical Modelling and Algorithms*. 7, 43-58. https://doi.org/10.1007/s10852-007-9074-5

Bramel, J., & Simchi-Levi, D. (1995). A location based heuristic for general routing problems. *Operations Research*, 43(4), 649-660. https://doi.org/10.1007/s10852-007-9074-5

Brimberg J., & Mladenovic, N. (1995). A variable neighbourhood algorithm for solving the continuous location-allocation problem. *GERAD*.

Chaves, A.A., Correa, F.D.A., & Lorena, L.A.N. (2007). Clustering search heuristic for the capacitated *p*-median problem. *Innovation in Hybrid Intelligent Systems*, 4, 136-143. https://doi.org/10.1007/s10852-007-9074-5

Diaz, J.A., & Fernandez, E. (2006). Hybrid scatter search and path relinking for the capacitated *p*-median problem. *European Journal of Operational Research,* 169(2), 570-585. https://doi.org/10.1016/j.ejor.2004.08.016

El Amrani, M., Benadaa, Y., & Gendron, B. (2016). Generalization of capacitated *p*-median location problem: Modeling and resolution. *3rd International Conference on Logistics Operations Management.* Fez, Morocco. https://doi.org/10.1109/GOL.2016.7731674

Fleszar, K., & Hindi, K.S. (2008). An effective VNS for the capacitated *p*-median problem. *European Journal of Operational Research.* 191(3), 612-622. https://doi.org/10.1016/j.ejor.2006.12.055

Garey, M.R., & Johnson, D.S. (1990). *Computers and intractability: A guide to the theory of NP-Completeness.* New York: Subs. of Scientific American, Inc.

Ghoseiri, K., & Ghannadpour, S.F. (2013). An efficient heuristic method for capacitated *p*-median problem. *International Journal of Management Science and Engineering Management*, 4(1). https://doi.org/10.1080/17509653.2009.10671064

Hansen, P, & Mladenovic, N. (1997). Variable neighborhood search for the *p*-median. *Location Science*, 5(4), 207-226. https://doi.org/10.1016/S0966-8349(98)00030-8

Hansen, P., Mladenovic, N., & Perez, J.A.M. (2008). Variable neighbourhood search: methods and applications. *4OR*, 6, 319-360. https://doi.org/10.1007/s10288-008-0089-1

Hansen, P., & Mladenovic, N. (2001). Variable neighborhood search: Principles and applications. *European Journal of Operational Research*, 130(3), 449-467. https://doi.org/10.1016/S0377-2217(00)00100-4

Irawan C.A., & Salhi, S. (2013). Solving large *p*-median problems by a multistage hybrid approach using demand points aggregation and variable neighbourhood search. *Journal of Global Optimization*, 63(3), https://doi.org/10.1007/s10898-013-0080-z

Irawan, C.A., Eskandarpour, M., Ouelhadj, D., & Jones, D. (2021). Simulation-based optimisation for stochastic maintenance routing in an offshore wind farm. *European Journal of Operational Research*, 289(3), 912-926. https://doi.org/10.1016/j.ejor.2019.08.032

Irawan, C.A., Imran, A., & Luis, M. (2017). Solving the bi-objective capacitated *p*-median problem with multilevel capacities using compromise programming and VNS. *International Transactions in Operational Research*, 27(1), 361-380. https://doi.org/10.1111/itor.12485

Irawan, C.A., Salhi, S., & Drezner, Z. (2016). Hybrid meta-heuristics with VNS and exact methods: application to large unconditional and conditional vertex *p*-centre problems. *Journal of Heuristics*, 22(4), 507-537. https://doi.org/10.1007/s10732-014-9277-7

Irawan, C.A., Salhi, S., & Scaparra, M.P. (2014). An adaptive multiphase approach for large unconditional and conditional *p*-median problems. *European Journal of Operational Research*, 237(2), 590-605. https://doi.org/10.1016/j.ejor.2014.01.050

Kariv, O., & Hakimi, S.L. (1979). An algorithmic approach to network location problems. I: The *p*-Centers. *SIAM Journal on Applied Mathematics*, 37(3), 513-538. https://doi.org/10.1137/0137040

Klein, G., & Aronson, J.E. (1991). Optimal clustering: A model and method. *Naval Research Logistic*, 38(3), 447-461. https://doi.org/10.1002/1520-6750(199106)38:3<447::AID-NAV3220380312>3.0.CO;2-0

Landa-Torres, I., Ser, J.D., Salcedo-Sanz, S., Gil-Lopez, S., Portilla-Figueras, J.A., & Alonso-Garrido, O. (2012). A comparative study of two hybrid grouping evolutionary techniques for the capacitated *P*-median problem. *Computers & Operations Research*, 39(9), 2214-2222. https://doi.org/10.1016/j.cor.2011.11.004

Li, X., Tian, P., & Leung, S.C.H. (2010). Vehicle routing problems with time windows and stochastic travel and service times: Models and algorithm. *International Journal of Production Economics*, 125(1), 137-145. https://doi.org/10.1016/j.ijpe.2010.01.013

Lorena, L.A.N., & Senne, L.F. (2004). A column generation approach to capacitated *p*-median problems. *Computers & Operations Research*, 31(6), 863-876. https://doi.org/10.1016/S0305-0548(03)00039-X

Maniezzo, V., Mingozzi, A., & Baldacci, R. (1998). A bionomic approach to the capacitated *p*-median problem. *Journal of Heuristics*. 4, 263-280. https://doi.org/10.1023/A:1009665717611

Mulvey, J.M., & Beck, M.P. (1984). Solving capacitated clustering problems. *European Journal of Operational Research*, 18(3), 339-348. https://doi.org/10.1016/0377-2217(84)90155-3

Osman, I.H., & Christofides, N. (1994). Capacitated clustering problems by hybrid simulated annealing and tabu search. *International Transactions in Operational Research*, 1(3), 317-336. https://doi.org/10.1016/0969-6016(94)90032-9

ReVelle, C.S., & Swain, R.W. (1970). Central facilities location. *Geographical Analysis*, 2(1), 30-42. https://doi.org/10.1111/j.1538-4632.1970.tb00142.x

Scheuerer, S., & Wendolsky, R. (2006). A scatter search heuristic for the capacitated clustering problem. *European Journal of Operational Research*. 169(2), 533-547. https://doi.org/10.1016/j.ejor.2004.08.014

Shen, Y., Xu, J., Wu, X., & Ni, Y. (2019). Modelling travel time distribution and its influence over stochastic vehicle scheduling. *Transport*, 34(2), 237-249. https://doi.org/10.3846/transport.2019.8940

Stefanello, F., Araujo, O., & Muller, F.M. (2014). Matheuristics for the capacitated *p*-median problem. *International Transactions in Operational Research*, 22(1). https://doi.org/10.1111/itor.12103

The World Meteorological Organization (2009). The state of greenhouse gases in the atmosphere using global observation through 2008. *WMO Gas Bulletin*. Technical Report. United Nation, Paris.

Wang, F., Lai, X., & Shi, N. (2011). A multi-objective optimization for green supply chain network design. *Decision Support Systems*, 51(2), 262-269. https://doi.org/10.1016/j.dss.2010.11.020

Yaghini, M., Karimi, M., & Rahbar, M. (2013). A hybrid metaheuristic approach for the capacitated *p*-median problem. *Applied Soft Computing*, 13(9), 3922-3930. https://doi.org/10.1016/j.asoc.2013.04.009

Journal of Industrial Engineering and Management, 2022 (www.jiem.org)