Soft Computing (2023) 27:12883–12893
https://doi.org/10.1007/s00500-022-06785-7

FOCUS

Unbiased decision making in location-routing problems with uncertain customer demands

Annarita De Maio[1,3] · **Roberto Musmanno**[1] · **Francesca Vocaturo**[2]

Accepted: 7 January 2022 / Published online: 10 February 2022

Abstract

In this paper, we focus on a location-routing problem (LRP) in the dairy industry. This problem consists of locating a cold storage warehouse, from which vehicles of limited capacity are dispatched to serve a given number of supermarkets with uncertain service requirements, and determining the order of supermarkets served by each vehicle. First, the LRP is solved by using a classical approach based on a deterministic model where the service requirements, i.e. customer demands, are defined through sample means. Second, we propose an indifference zone approach to the LRP. The indifference zone procedures are specific ranking and selection methods aimed at selecting the best option from a set of alternative configurations. In particular, they attempt to guarantee the probability of correct choice, while minimising the computational effort. The numerical results presented in the paper highlight the risk of biased decision making when mere sample means are used in a deterministic model. In addition, they show the effectiveness of indifference zone approaches to the dairy products distribution activity.

Keywords Location-routing · Demand uncertainty · Indifference zone · Confidence interval

1 Introduction

Managing supply chains represents a major challenge due to the complexity of interactions among the system components, incomplete information available for decision making, the presence of multiple decision makers pursuing different objectives and the dynamics of non-stationary conditions (von Lanzenauer and Pilz-Glombik 2002). Coordinating information and material flows in an effective way is fundamental to supply chain management. The coordination can be realised by using optimisation approaches to support decisions at strategic, tactical and operational levels. The strategic level designs the logistics network, including prescribing facility locations, production technologies and plant capacities. The tactical level specifies material flow management policies. Finally, the operational level schedules operations to assure in-time delivery of final products to customers (Schmidt and Wilhelm 2000). Optimisation technologies underpin the development of decision support systems in a wide variety of applications (Barbosa Póvoa et al. 2017). In particular, a large number of logistics problems in supply chain management can be faced through the formulation and consequent solution of mathematical models, characterised by a set of decision variables, of continuous or discrete type, in function of which the outcome to optimise is defined. In addition, there exists a set of constraints to satisfy, which are generally equations or inequalities of the decision variables. In most industrial sectors, real-world problems usually lead to large-scale models. We refer the reader to the paper of Grossmann (2014) for a comprehensive survey on enterprise-wide optimisation which involves optimising the supply, manufacturing and distribution activities of a company.

Communicated by Dario Pacciarelli.

✉ Roberto Musmanno
roberto.musmanno@unical.it

Annarita De Maio
annarita.demaio@unical.it

Francesca Vocaturo
vocaturo@unical.it

1 Department of Mechanical, Energy and Management Engineering - DIMEG, University of Calabria, Via P. Bucci, Cubo 41C, 87036 Arcavacata di Rende, CS, Italy

2 Department of Economics, Statistics and Finance “Giovanni Anania” - DESF, University of Calabria, Via P. Bucci, Cubo 0C, 87036 Arcavacata di Rende, CS, Italy

3 Present Address: Department of Economics, Statistics and Finance “Giovanni Anania” - DESF, University of Calabria, Via P. Bucci, Cubo 0C, 87036 Arcavacata di Rende, CS, Italy

Here we deal with the distribution activity and focus on the nature of the input parameters used into the mathematical models that frequently are assumed of deterministic type. This assumption may lead to unpleasant results from an economic prospective. The case that will be examined in this paper is rather emblematic to demonstrate empirically this statement. Specifically, we refer to a logistics problem known as location-routing problem (LRP) and highlight the importance of considering uncertainty adequately. The LRP aims at determining the best facility locations among a set of potential facility sites to minimise a total logistics cost which also includes the transport cost paid to serve the customers from the activated facilities (Mara et al. 2021). This problem arises in case of less-than-full-load shipments, when the same vehicle can serve multiple customers; as a result, the transport cost depends on the order in which the customers to be served are visited and, consequently, the optimal facility locations depend on the vehicle routes. The location decision is usually made at a strategic level, while vehicle routes have to be defined at an operational level. However, location and routing decisions are interdependent and many studies in the scientific literature confirm that the overall logistics cost can be significantly larger than the minimum if the location of facilities and the design of routes are tackled separately.

In this paper we consider the case of an LRP solved by a dairy firm, where the customer demand values are characterised by high variability. We point out that these values impact on the vehicle capacity and, consequently, on the number and kind of vehicle routes. In detail, our location problem involves the evaluation of three different alternatives (i.e. potential sites) where a cold storage warehouse can be located. Starting from the selected site, some vehicles have to serve 12 supermarkets distributed in the surrounding region. The operational part of the problem deals with the determination of the daily vehicle routes; to this end, the company can use its own fleet and, if necessary, activate a third-party provider to cover demand peaks. For the problem, we propose two different approaches. The first one fits into the area of deterministic optimisation and is based on a mathematical model where the values of the demands are replaced by sample means. The second approach fits into the area of optimisation under uncertainty and is based on ranking and selection. According to Fu and Henderson (2017), the ranking and selection techniques provide "statistical guarantees" on the chosen solution and lead to unbiased decisions. For an introduction to ranking and selection we refer the reader to the expository paper by Gibbons et al. (1979) where the philosophy of the indifference zone strategy used in our paper is described in detail. The computational experiments show that the two approaches lead to significant differences in terms of outcome and point out the necessity of tackling logistics problems within the framework of optimisation under uncertainty whenever the customer demand values are characterised by high variability.

The use of a ranking and selection technique, namely a non-traditional approach to this class of problems, makes our contribution different to others that focus on LRPs under uncertainty. We recognise the importance of traditional modelling frameworks like stochastic, robust and fuzzy optimisation for LRPs, but the approaches based on these frameworks are generally complex and not immediately applicable. Instead, we propose a simple indifference zone method that can work well when perishable products have to be delivered. In effect, frequently in this context the routes are enumerable since they are relatively few and small, both in terms of kilometres and customers to be served. In particular, our approach can provide, without a high computational effort, effective and easy-to-be-accepted answers to managers in dairy industry.

The remainder of the paper is organised as follows. Section 2 provides a literature review. Section 3 illustrates the formulation and the two proposed solution approaches. Section 4 describes the case study, including the input data, the computational experiments, and the numerical results. Finally, Sect. 5 highlights general conclusions and future research directions.

2 Literature review

The most related studies are reviewed in the next sections.

2.1 Studies on optimisation under uncertainty

In recent decades, the ever-increasing complexity and difficulty of real-world problems resulted in the need for more reliable optimisation techniques, especially metaheuristic algorithms (see Abualigah et al. 2021a, c for very recent metaheuristic schemes). The results presented in the scientific literature are scattered due to the difficulty of finding an effective modelling framework for real-world problems that differ greatly from one another. A further complication arises whenever uncertainty is involved in the decision-making process. The problems of optimisation under uncertainty are characterised by the necessity of making decisions without knowing what their full effects will be (Sahinidis 2004). The main paradigms dealing with uncertain data in problem-solving situations are stochastic programming and robust optimisation. According to De Maio et al. (2021), stochastic programming deals with cases in which the decision maker is interested in building a solution that is good on average. This is highly relevant for repeated decision-making processes like portfolio investment. Instead, robust optimisation is appropriate for processes where one has to decide once and for all (like network design); in addition, it is frequently applied in decisions making with high impact, where "bad surprises" should be avoided. In the area of logistics and supply chain management both paradigms are frequently used (see, for example, Beraldi et al. 2010; Laporte et al. 2010; Renaud et al. 2017; Hemmati Golsefidi and Akbari Jokar 2020; Sangaiah et al. 2020).

Uncertainty in optimisation also refers to vagueness that arises in fuzzy environments. Although fuzzy logic has multiple connotations, the common starting point is the classic principle of bivalence. It states that any declarative sentence has only two possible truth values: true and false. Fuzzy logic considers not only these two values, but allows for additional ones. From this perspective, expressions like "roughly", "approximately", "around" are usual. Practically, fuzzy models offer the opportunity to model subjective imaginations of the decision maker as precisely as a decision maker will be able to describe them. An example of application of fuzzy models to logistics problems is provided by Zhang et al. (2020). Other modelling frameworks have been proposed in the literature for optimisation under uncertainty like simulation and simulation-based optimisation (Bierlaire 2015; Legato et al. 2021). More generally, we can refer to sampling-based paradigms that do not necessarily use a simulator. Along with them, a variety of methods have been developed and used successfully in many applications. For instance, Fu and Henderson (2017) mention the selection methods as a collection of techniques for identifying the best from a (small) set of alternative solutions. Specifically, these techniques determine how many samples need to be collected from each alternative and then which alternative should be selected as the best based on the sample information. Among the selection methods, there are the ranking and selection techniques that provide "statistical guarantees" on the chosen solution. For a comprehensive overview, we refer the reader to the recent paper by Hong et al. (2021) in which theoretical achievements on ranking and selection and practical applications in the past 20 years are discussed. Hong et al. (2021) also explain the philosophy of the indifference zone approach to ranking and selection. This approach has a rich literature. An extensive amount of research has been published since the indifference zone concept was introduced by Bechhofer (1954). It includes the recent study of Yoon and Bekker (2019).

2.2 Studies on location-routing

As mentioned in Sect. 1, the LRP is a logistics problem aimed at optimising two different decisional levels in an integrated way: the strategical decision on facility location and the operational decision on routing. Watson-Gandy and Dohrn (1973) were the first to clearly consider the cost of visiting customers while locating depots. Specifically, in order to represent the multi-drop nature of the journeys performed by the drivers, these authors used a nonlinear function of distance rather than straight-line distances. However, the discussion on the interdependency of location and routing decisions has already started earlier (see, for example, von Boventer 1961). Meanwhile, many managers at that time were realising the drawbacks on separating these decisions, as affirmed by Rand (1976) "Many practitioners are aware of the danger of suboptimising by separating depot location from vehicle routing". Then, when the renowned work by Salhi and Rand (1989) was published, the advantage of considering location and routing problems simultaneously was fully realised. The discussion about the LRP has gained popularity in the subsequent years and led to a huge quantity of scientific works. Many studies focus on practical applications of LRPs, e.g. delivery services (Perl and Daskin 1984; Bruns et al. 2000; Wasner and Zäpfel 2004), military equipment location (Murty and Djang 1999), waste collection (Kulcar 1996) and postbox location (Labbé and Laporte 1986).

The literature also proposes many surveys dedicated to this topic. For instance, Nagy and Salhi (2007) present a classification of the LRPs and describe exact and heuristic methods for the basic version and its extensions. Prodhon and Prins (2014) and Drexl and Schneider (2015) introduce a more precise classification and consider further variants: multi-echelon, mobile depots, trucks and trailers, multiple objective, inventory management and multi-period problems. Hassanzadeh et al. (2009), Schneider and Drexl (2017), and Mara et al. (2021) present other surveys. In particular, Mara et al. (2021) perform a review of recent LRP research from 222 journal papers published from 2014 to 2019.

Since both location and routing problems are NP-hard under most scenarios, accordingly, their combination leads mostly to an NP-hard problem where the solution space increases exponentially with the size. Then, large-sized instances can hardly be solved by exact approaches (exception being the round-trip location problem, where routes only have to service pairs of nodes, to which polynomial time algorithms have already been proposed as pointed out by Lopes et al. (2013)). To date, the metaheuristic algorithms represent the most popular option to solve an LRP model. Most studies examined by Mara et al. (2021) (i.e. 67.12%) are based on metaheuristic schemes. In particular, Mara et al. (2021) specify that the variants of simulated annealing and genetic algorithm are favoured for single-objective LRP, while non-dominated sorting genetic algorithm II and multi-objective particle swarm optimisation are the preferred paradigms for multi-objective LRP. An interesting comparison among algorithms based on different paradigms is proposed by Lopes et al. (2013) with respect to three widely used sets of benchmark instances; for these algorithms the authors report the average computing time and the average percentage gap between the obtained results and the best-known lower bound.

It is worth noting that most studies in the scientific literature are based on deterministic modelling frameworks. In contrast, the operational part of the LRP (i.e. routing) is subject to a high variability when the input parameters change. Moreover, since the location decision has an impact on the medium–long planning horizon, it is difficult to predict these parameters deterministically in an accurate way. Therefore, there is a growing interest in formulating LRPs through the modelling frameworks described in Sect. 2.1. Uncertain parameters (demand levels, presence of customers, travel times, etc.) can be represented as random variables, whose probability distributions are evaluated on the basis of historical data.

In the paper of Albareda-Sambola et al. (2007) an uncertain LRP is formulated as a stochastic program with recourse and solved through a two-stage approach: in the first stage both the selected depots and the routes are defined, in the second stage a recourse action is applied to adapt the routes to the set of customers requiring service, once they are known. Klibi et al. (2010) deal with an LRP characterised by multiple transport options, multiple demand period and stochastic demands. Their problem is also formulated as a stochastic program with recourse and solved by a hierarchical heuristic approach. Other contributions focusing on uncertain parameters consider: probabilistic travel times (Ghaffari-Nasab et al. 2013), fuzzy travel times and time windows (Zarandi et al. 2011), fuzzy demands (Golozari et al. 2013), stochastic facility availability (Hassan-Pour et al. 2009) and stochastic inventory (Ahmadi-Javid and Azad 2010). Examples of LRPs under uncertainty in arc routing contexts are reported in the very recent survey of De Maio et al. (2021). For instance, Mirzaei-Khafri et al. (2020) propose a robust optimisation model for a location-arc routing problem in which the demand of each road is associated with a value that belongs to a bounded uncertainty set.

Finally, we point out that the theoretical analyses proposed in the scientific literature have been tested on instances not derived from real-world cases, with few exceptions. An interesting practical application that brings uncertainty on input parameters for an LRP is proposed by Chan et al. (2001). Specifically, the authors refer to a medical evacuation case study of the US Air Force.

3 The location-routing problem

The LRP considered in this paper deals with a single facility to be selected from a list I of potential sites. The facility serves a set V of customers. Each customer $j \in V$ is characterised by a demand d_j, whose value is not known and can only be estimated from a sample of finite size. Let $\bar{d}_j$ be the corresponding sample mean.

For the sake of simplicity, assume that the costs, which are dependent on the selected location $i \in I$, are only those related to vehicle routing. Let K_i be the set of *plausible* routes, each of which starting from i, serving a subset of customers respecting all the deterministic constraints involved in the problem (i.e. time duration of the route). Let a_{kj} be a binary constant, used to identify if a customer $j \in V$ is served by route $k \in K_i$ (in this case $a_{kj} = 1$) or not ($a_{kj} = 0$). Let c_k be the cost associated with route $k \in K_i$. Each route $k \in K_i$ is assigned to a vehicle of capacity q and there are n vehicles of this type. A plausible route becomes feasible if the capacity constraint is satisfied for the sample values of the demand associated with the customers served by the route. If a customer $j \in V$ cannot be served by the available fleet of vehicle, it is possible to outsource the service delivery to a third-party provider at a cost equal to f_j. Let y_k be a binary decision variable equal to 1 if the corresponding route $k \in K_i$ is selected, 0 otherwise. Let x_j be a binary decision variable equal to 1 if the corresponding customer $j \in V$ is served by the third-party provider, 0 otherwise. By considering the sample mean $\bar{d}_j$ of the demand d_j for each customer $j \in V$, it is possible to consider a deterministic version of the LRP by solving for each site location $i \in I$ the following binary programming model:

$$\text{Minimise} \sum_{k \in K_i} c_k y_k + \sum_{j \in V} f_j x_j \tag{1}$$

subject to

$$\sum_{k \in K_i} a_{kj} y_k + x_j \geq 1, \; j \in V \tag{2}$$

$$\left(q - \sum_{j \in V} a_{kj} \bar{d}_j \right) y_k \geq 0, \; k \in K_i \tag{3}$$

$$\sum_{k \in K_i} y_k \leq n \tag{4}$$

$$y_k \in \{0, 1\}, \; k \in K_i \tag{5}$$

$$x_j \in \{0, 1\}, \; j \in V. \tag{6}$$

Constraints (2) ensure that each customer is served. Constraints (3) ensure the feasibility of the routes in terms of vehicle capacity. In particular, if $(q - \sum_{j \in V} a_{kj} \bar{d}_j) < 0$ for some $k \in K_i$, then the vehicle capacity constraint is violated and, consequently, y_k should be 0. Constraint (4) limits to n the number of identical vehicles that can be used.

A possible drawback of formulation (1)–(6) is that the number of variables $y_k, k \in K_i$, may be very large especially for weakly constrained problems. However, there exist applications in which $|K_i|$ is relatively small, for each $i \in I$, like the case examined in this paper. For example, if the number of customers along a route is at most three, then $|K_i| = O\left(\binom{|V|}{3} + \binom{|V|}{2} + \binom{|V|}{1}\right) = O(|V|^3)$, for each $i \in I$.

The optimal objective function value of formulation (1)–(6) corresponds to an average routing cost indicated by $\bar{z}_i$; this value is associated with the choice of opening the facility in site $i \in I$. By solving the same model for each possible facility location within the set of potential sites, it is possible to determine, in a deterministic way, the best site i^* as

$$i^* = \arg\min_{i \in I} \bar{z}_i.$$

However, this approach is too simplistic in practical applications. It is worth observing that considering only the sample mean of the demand values flattens the variability in the demand realisations. For instance, in model (1)–(6) a generic customer is expected to be always served by a company vehicle or a third-party provider whereas, in practice, both situations can happen in different time periods on account of the effective realisations of its demand over the time.

Actually, the demand values are subject to fluctuations that cannot be effectively represented by deterministic models. The demand is naturally affected by uncertainty and this aspect should be incorporated in the LRP solution process. Then, an alternative approach is based on the following considerations. Since the customer demand can be represented by a vector $\mathbf{\Theta}$ of random variables, the outcome for each alternative site $i \in I$ can be seen as a function g_i of the vector $\mathbf{\Theta}$ corresponding to the unknown data. Hence, the outcome is a random variable itself, denoted as Z_i:

$$Z_i = g_i(\mathbf{\Theta}), i \in I.$$

In this case, the best site i^* is defined as

$$i^* = \arg\min_{i \in I} E_{\mathbf{\Theta}}[Z_i],$$

and corresponds to the location characterised by the least expected cost.

Since the probability function of $\mathbf{\Theta}$ is not known, but only m observations of $\mathbf{\Theta}$ are known, the choice of the site associated with the least expected cost is complex. In effect, the value $E_{\mathbf{\Theta}}[Z_i]$ corresponding to site $i \in I$ is not known and can only be estimated from the corresponding outcome samples $z_{i1}, z_{i2}, \ldots, z_{im}$. A point estimate of $E_{\mathbf{\Theta}}[Z_i]$ is provided by the sample mean

$$\overline{Z}_i = \frac{1}{m}\sum_{k=1}^{m} z_{ik}. \tag{7}$$

In addition, it is also required computing the confidence interval for the expected outcome, i.e. the interval in which $E_{\mathbf{\Theta}}[Z_i]$ falls with a prescribed confidence level $(1-\alpha)$. Indeed, a point estimate does not necessarily coincide with the real expected value, while a confidence interval is more reliable. The confidence interval at $(1-\alpha)$ level of $E_{\mathbf{\Theta}}[Z_i]$, $i \in I$, is defined as

$$Pr\left(\overline{Z}_i - t_{\alpha/2,m-1}\frac{S_i}{\sqrt{m}} \le E_{\mathbf{\Theta}}[Z_i] \le \overline{Z}_i + t_{\alpha/2,m-1}\frac{S_i}{\sqrt{m}}\right) = 1-\alpha,$$

or, equivalently,

$$Pr\left(E_{\mathbf{\Theta}}[Z_i] < \overline{Z}_i - t_{\alpha/2,m-1}\frac{S_i}{\sqrt{m}} \text{ or } E_{\mathbf{\Theta}}[Z_i] > \overline{Z}_i + t_{\alpha/2,m-1}\frac{S_i}{\sqrt{m}}\right) = \alpha,$$

where $t_{\alpha/2,m-1}$ is the quantile of order (1-$\alpha/2$) of the Student's t-distribution with $m-1$ degrees of freedom and S_i is the sample standard deviation

$$S_i = \sqrt{\frac{1}{m-1}\sum_{k=1}^{m}(z_{ik} - \overline{Z}_i)^2}. \tag{8}$$

Therefore, the confidence interval $(1-\alpha)$ of $E_{\mathbf{\Theta}}[Z_i]$, $i \in I$, is

$$\left[\overline{Z}_i - t_{\alpha/2,m-1}\frac{S_i}{\sqrt{m}}, \overline{Z}_i + t_{\alpha/2,m-1}\frac{S_i}{\sqrt{m}}\right].$$

When the confidence intervals overlap, selecting the best solution should correspond to select the alternative i^* in I of least expected cost with a certain confidence level $(1-\alpha)$. To this end, an indifference zone method like the Rinott's procedure can be used (Hong et al. 2021). The difficulty in applying the procedure depends on how much i^* is better than the other alternatives. If i^* is considerably better, then it should be easy to identify. In contrast, if there is at least one site very close to i^*, it may be difficult to select the correct solution using sample data. However, if two or more alternatives are so close to being the best, it will be irrelevant which one is selected as the best. The Rinott's procedure, as well as the other indifference zone methods, exploits this idea: it is assumed that the decision maker is indifferent to selecting an alternative that does not yield the least cost if its expected value falls into an indifference zone defined by parameter δ. Given a sample size m of the demand vector, the Rinott's procedure allows to select the alternative with the least sample mean as the best (or an alternative with an outcome lying within δ of the best one) with probability $(1-\alpha)$; m observations could be insufficient to evaluate an alternative and extra observations could be required to draw a conclusion. The details can be found in Algorithm 1. Specifically, given the input values α, $|I|$, m, and δ, RINOTT(α, $|I|$, m, δ, $\mathbf{b}$, i^*) returns i^* and a vector $\mathbf{b}$ with $|I|$ components. The i^{th} component of $\mathbf{b}$, i.e. b_i, corresponds to the number of samples (observations) used to evaluate alternative i, $i \in I$. The procedure initially determines the Rinott's parameter r by solving the following equation:

$$\int_0^\infty \left[\int_0^\infty \Phi\left(\frac{r}{[(m-1)(1/x+1/y)]^{1/2}}\right) f(x)dx\right]^{|I|-1} f(y)dy = (1-\alpha), \tag{9}$$

where f represents the density of a χ^2 random variable with $m-1$ degrees of freedom and $\Phi(x)$ represents the cumulative distribution function of the standard normal distribution. For more details, see Wilcox (1984).

Algorithm 1 RINOTT(α, $|I|$, m, δ, $\mathbf{b}$, i^*)

```
Compute the Rinott's constant r by using (9);
for i ∈ I do
    Compute the sample mean Z̄_i and the sample standard deviation S_i on the basis of the m observations available for each alternative i by using (7) and (8), respectively;
    Set γ = ⌈(r S_i / δ)^2⌉ (number of observations needed for alternative i);
    if γ > m then
        Take γ − m more observations for alternative i;
        Set b_i = γ;
        Set Z̄_i = (1/b_i) Σ_{k=1}^{b_i} z_ik;
        Set S_i = sqrt( (1/(b_i − 1)) Σ_{k=1}^{b_i} (z_ik − Z̄_i)^2 );
    else
        Set b_i = m;
    end if
end for
Select the alternative of least cost, i.e. i* = arg min_{i∈I} Z̄_i;
return b, i*;
```

4 The case study

The case study illustrated in this section deals with the sector of the dairy products. The farm is headquartered in France and the problem consists of locating a cold storage warehouse. It will be supplied directly by the main production plant and used to serve daily 12 supermarkets of different sizes located in the region of Auvergne (see Table 1). The daily demand of the dairy products requested by each supermarket can be estimated by using historical data; the minimum, the maximum and the most probable values (measured in kilograms) are reported in Table 1. It is worth observing that the data entries show a constant trend with a weekly seasonality and a significant ($\pm$20%) random component. To have a more precise idea of the variability of the demand pattern, Fig. 1 is helpful; it depicts the plot of the corresponding time series for each supermarket in a time horizon of 21 days (three weeks). The values of the complete time series are available from the corresponding author upon request.

Three different potential sites are considered to host the cold storage warehouse (i.e. $I = \{1, 2, 3\}$). They are located in the towns of Ennezat, Vic-le-Comte and Saint Ours, respectively (see Fig. 2 and Table 2). The warehouse is also used as a depot for the fleet of four identical refrigerator vans, each of which has a capacity of 1,100 kilograms.

The company produces and distributes the following types of finished products: fresh bottled milk, long shelf-life bottled milk, ricotta, mozzarella, long shelf-life cheese, fresh cheese. In the distribution activity, the company utilises three different forms of packaging: plastic boxes, cartons and palletized unit loads (i.e. pallets of one metre long and wide). In order to standardise the three types of unit loads and determinate the vehicle load capacity, the equivalent unit load (EUL) is introduced (pallet with an average weight of 300 kilograms). Consequently, the load capacity of each vehicle is 3.66 EULs. The weights of each plastic box and carton are 17.80 kilograms and 24.19 kilograms on average, equal to 0.06 EULs and 0.08 EULs, respectively. We observe that 3.66 EULs generates a volume always less than the van volume; therefore the only physical characteristic to consider in the van packing is the weight.

A set K_i of plausible routes can be preliminarily identified for each potential site $i \in I$. Each route starts at the selected depot, serves at most three supermarkets and ends at the depot. This means that $|K_i| = \binom{12}{1} + \binom{12}{2} + \binom{12}{3} = 298$, $i \in I$. It is worth observing that each route satisfies the van capacity constraint, with respect to the most probable values of the daily demand of dairy products associated with the supermarkets served by the route. The cost of each route can be computed by considering a unit cost of 0.25 euros for kilometre multiplied by the length of the route. Note that only the minimum-length route is considered for each subset of customers. The kilometric lengths of all plausible routes are data available from the corresponding author upon request. The company can use a third-party service provider that serves, if necessary, the unserviced customers through direct routes, at a cost that is composed of a fixed term of 30 euros, plus a variable part which is proportional to the distance between the potential site and the supermarket, with a unit cost equal to 1.23 euros for kilometre for trips longer than 50 kilometres and to 1.43 euros for kilometre for trips with length less than or equal to 50 kilometres.

First, the objective function value corresponding to each alternative $i \in I$ was determined by using a deterministic approach, i.e. considering $\overline{d}_j$ equal to the average demand in kilograms for each supermarket j (see the last column of Table 1). Specifically, model (1)–(6) was solved for each of the three candidate sites to host the cold storage warehouse. We used GAMS 24.7.4 (GAMS Development Corporation) as algebraic modelling system, with CPLEX 12.6 (IBM Cor-

Table 1 Geographical coordinates and demand characteristics of each supermarket

Supermarket Id	Latitude	Longitude	Minimum demand (kg)	Maximum demand (kg)	Average demand (kg)
1	45.9167820762827	2.6682355704796	156	381	255
2	45.8766406890607	3.5086896171367	216	525	355
3	45.7713709583362	3.2093121952883	264	639	426
4	45.5767197363509	3.7398078237765	264	639	429
5	45.9290511483901	3.1183272949744	210	516	344
6	45.8302793429363	3.1316894350502	279	672	448
7	45.5617051790017	2.9229492143118	228	549	369
8	45.8379339374907	3.0135864154219	144	360	240
9	45.8034799706812	3.0685180524584	156	378	252
10	45.7824142728244	3.5436767128234	150	366	243
11	45.6174443249854	2.9504150328301	144	360	239
12	46.2055287600722	3.4554948273149	222	546	366

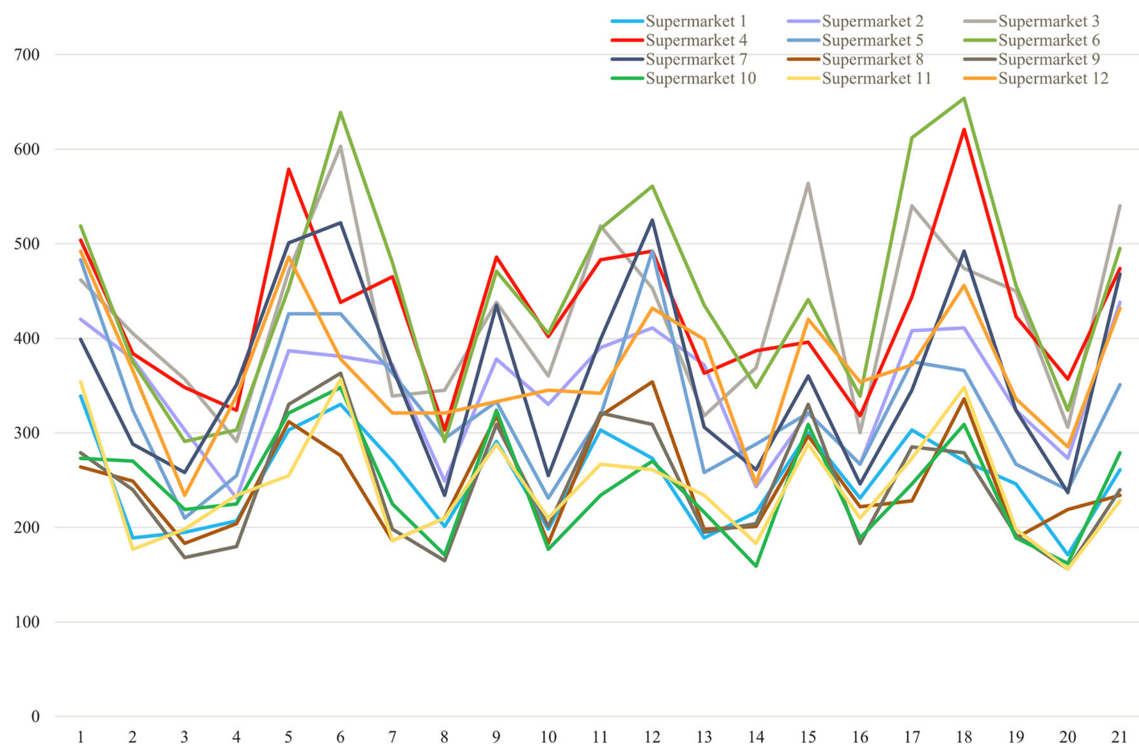

Fig. 1 Daily demand (in kilograms) of the dairy products required by the 12 supermarkets

poration) as solver. All the experiments were carried out on a PC Intel Core i7 (2.3 GHz) with 12 GB of RAM. The computational times were quite short. Therefore, they are not reported here.

The average daily transport costs (in euros), obtained by solving to optimality the model (1)–(6) for each of the three alternatives correspond to 131.93, 139.73 and 144.25 respectively for Ennezat, Vic-le-Comte and Saint Ours. Note that the costs do not include the fixed and the variable facility costs which are constant. The best solution corresponds, therefore, to the site located in Ennezat. The daily variable transport cost corresponds to the sum of the costs associated with the activated routes. We point out that no external service is selected.

However, this solution does not capture the variability of the daily demand for dairy products. For this reason, the Rinott's procedure was introduced. Consequently, a confidence interval on the expected daily cost was also computed, with a confidence level $(1 - \alpha) = 0.95$ and an indifference zone of 6 euros. The optimisation model was solved $m = 480$ times for each potential town selected as depot, on the basis of the different realisations of the daily demand. It is worth observing that, considering $\alpha = 0.05$, $m = 480$

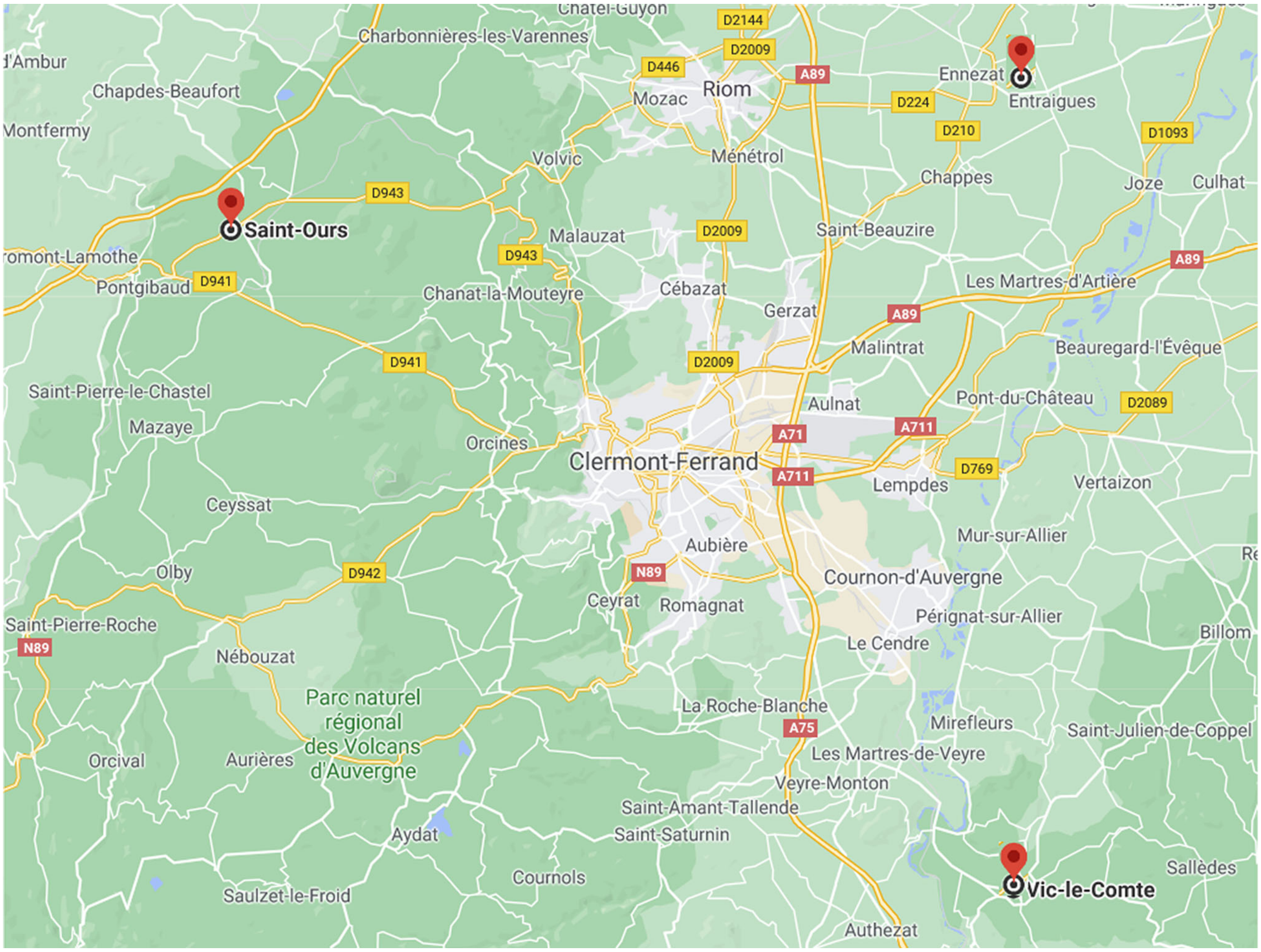

Fig. 2 French region where the three potential facilities are located (image taken from Googlemaps®)

Table 2 Geographical coordinates of the three potential sites

Potential site	Latitude	Longitude
Ennezat	45.9219755703321	3.2226196167101
Vic-le-Comte	45.6712672424955	3.2509691228137
Saint Ours	45.8740299569840	2.8867817259801

and $|I| = 3$, the Rinott's constant r is equal to 2.7704. Therefore, the following values were computed: $\lceil (\frac{rS_1}{\delta})^2 \rceil = \lceil (\frac{2.7704 \times 37.80}{6})^2 \rceil = 305$, $\lceil (\frac{rS_2}{\delta})^2 \rceil = \lceil (\frac{2.7704 \times 39.05}{6})^2 \rceil = 326$, and $\lceil (\frac{rS_3}{\delta})^2 \rceil = \lceil (\frac{2.7704 \times 40.32}{6})^2 \rceil = 347$. Since they were less than m, no additional observation on the daily demand was required to draw a conclusion about the site selection.

The results obtained by applying the second approach, based on the indifference zone philosophy, are summarised in Table 3. Because of the variability of the daily demand, the van capacity constraint is not satisfied for some realisations and the fleet of vans is not sufficient to serve all supermarkets. In this case, the best solution corresponds to the site located in Vic-le-Comte. This solution is different from the one obtained by only considering the average daily demands. Although Vic-le-Comte has the largest service provider activation percentage, it remains the most competitive solution in terms of daily transport cost.

It is worth noting that the scenario analysis allows the logistics manager to focus on a critical aspect of the distribution activity, that is, the necessity to outsource part of the delivery service. We observe that in the 43.54% of the cases (i.e. 209 over the 480 tests) it is necessary to use a third-party provider for serving at least one supermarket. Note that 53 of the 209 cases involve external service for two supermarkets, and only in one case it is necessary to serve simultaneously three supermarkets. This aspect does not arise in the first case and highlights the need to manage demand peaks in a more structured way. On the basis of the high percentage of deliveries given in outsourcing, the logistics manager could assess cost-convenient contracts with a third-party provider to avoid excessive costs or stock-outs in certain circumstances. Alter-

Table 3 Summary of the results obtained by considering the 480 scenarios

	City		
	Ennezat	Vic-le-Comte	Saint Ours
Sample mean (€/day)	164.92	162.36	177.90
Sample standard deviation (€/day)	37.80	39.05	40.32
Confidence interval (€/day)	(161.53; 168.31)	(158.85; 165.86)	(174.28; 181.51)
Third-party provider activation (%)	43.54	43.54	42.92

natively (or additionally), the manager could carry out an investment evaluation for the long-term horizon, in order to expand the internal fleet of vehicles and avoid (or reduce) the cost of outsourcing.

In summary, in the location-routing context, the application of a ranking and selection method can lead to several advantages from a managerial prospective. First, it offers the opportunity to make specific and robust decisions by considering uncertainty and avoiding day-to-day emergencies (e.g. the need to frequently contact a service provider). Second, the application of a ranking and selection method is generally associated with a short computational time when the number of alternative solutions is not large. Finally, it leads to easy-to-be-accepted decisions for a non-expert audience and provides "statistical guarantees" on the solution. More generally, a ranking and selection approach can be successfully applied to other logistics problems in which strategical or tactical decisions have to be made also considering operational issues (inventory-routing problems, production-inventory-routing problems, etc.).

5 Conclusions

In this paper, we have considered a location-routing problem (LRP) with the aim of locating a cold storage warehouse in the dairy industry. From this warehouse, some refrigerator vans of limited capacity are dispatched to serve a set of local supermarkets. The demands associated with the supermarkets are characterised by high variability. In this case, the approaches based on paradigms of optimisation under uncertainty seem to be more adequate than the classical approaches based on deterministic models where sample means are used as input parameters. Specifically, we have proposed for the LRP a ranking and selection method exploiting the concept of indifference zone. The method is based on an indifference zone parameter, which refers to the smallest mean difference worth detecting. Several alternatives may have means that fall into the indifference zone, i.e. be good alternatives. According to the definition of indifference zone, the decision maker should be indifferent if one of these good alternatives is selected as the best (Hong et al. 2021). The solution obtained by the ranking and selection method is different with respect to the solution of the deterministic counterpart. Our computational experience has confirmed that the ranking and selection method quickly provides responses that are unbiased and easy to be accepted for the managers. In addition, it meets the need of a lean and fast computational analysis and can be adopted to effectively manage unexpected demand peaks in the supply chain.

It is possible to imagine future developments for the work presented in this paper. As mentioned in Sect. 2, there exist alternative ways to tackle uncertainty in the LRP. For instance, there are some contexts in which the size of the instances is such that the ranking and selection method proposed in this paper would result ineffective. This may arise for non-perishable products that do not require daily deliveries, especially, early in the morning. When high-quality routes for the specific problem are not easily enumerable, other methods may result more effective. For instance, it is possible to design a simulation–optimisation framework for LRPs under uncertainty in supply chains of different nature. These frameworks frequently incorporate ranking and selection procedures to evaluate the correct number of simulation runs (see, for example, Laganà et al. 2006; Ghiani et al. 2007). An alternative future research direction focuses on analysing LRPs in different supply chains by using stochastic programming or robust optimisation.

Finally, further research may focus on the opportunity of delivering dairy products within a smart city scheme (Abualigah et al. 2021b; Evtodieva et al. 2020) or, alternatively, modelling the problem as a green LRP (Dukkanci et al. 2019; Wang et al. 2020).

Author Contributions All authors proposed the model and designed the decision support method. R. Musmanno and F. Vocaturo reviewed the scientific literature. A. De Maio implemented the algorithms and collected the computational results. All authors discussed and wrote the paper together. The contribution of all authors to the paper is equal.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

Abualigah L, Diabat A, Mirjalili S, Abd Elaziz M, Gandomi AH (2021a) The arithmetic optimization algorithm. Comput Methods Appl Mech Eng 376:113609

Abualigah L, Diabat A, Sumari P, Gandomi AH (2021b) Applications, deployments, and integration of Internet of Drones (IoD): a review. IEEE Sens J 21(22):25532–25546

Abualigah L, Yousri D, Abd Elaziz M, Ewees AA, Al-qaness MAA, Gandomi AH (2021c) Aquila optimizer: a novel meta-heuristic optimization algorithm. Comput Ind Eng 157:107250

Ahmadi-Javid A, Azad N (2010) Incorporating location, routing and inventory decisions in supply chain network design. Transp Res Part E: Logist Transp Rev 46(5):582–597

Albareda-Sambola M, Fernández E, Laporte G (2007) Heuristic and lower bound for a stochastic location-routing problem. Eur J Oper Res 179:940–955

Barbosa Póvoa AP, Corominas A, de Miranda JL (eds) (2017) Optimization and Decision Support Systems for Supply Chains. Springer, Berlin

Bechhofer RE (1954) A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann Math Stat 25(1):16–39

Beraldi P, Ghiani G, Musmanno R, Vocaturo F (2010) Efficient neighborhood search for the probabilistic multi-vehicle pickup and delivery problem. Asia-Pac J Oper Res 27(3):301–314

Bierlaire M (2015) Simulation and optimization: a short review. Transp Res Part C: Emerg Technol 55:4–13

Bruns A, Klose A, Stähly P (2000) Restructuring of Swiss parcel delivery services. OR Spectrum 22:285–302

Chan Y, Carter WB, Burnes MD (2001) A multiple-depot, multiple-vehicle, location-routing problem with stochastically processed demands. Comput Oper Res 28(8):803–826

De Maio A, Laganà D, Musmanno R, Vocaturo F (2021) Arc routing under uncertainty: introduction and literature review. Comput Oper Res 135:105442

Drexl M, Schneider M (2015) A survey of variants and extensions of the location-routing problem. Eur J Oper Res 241(2):283–308

Dukkanci O, Kara BY, Bektaş T (2019) The green location-routing problem. Comput Oper Res 105:187–202

Evtodieva T, Chernova D, Ivanova N, Wirth J (2020) The Internet of Things: possibilities of application in intelligent supply chain management. In: Ashmarina S, Mesquita A, Vochozka M (eds) Digital transformation of the economy: challenges, trends and new opportunities. Advances in intelligent systems and computing, vol 908. Springer, Berlin, pp 395–403

Fu MC, Henderson SG (2017) History of seeking better solutions, aka simulation optimization. In: Chan WKV, D'Ambrogio A, Zacharewicz G, Mustafee N, Wainer G, Page E (eds) Proceedings of the 2017 Winter Simulation Conference. IEEE, pp 131–157

Ghaffari-Nasab N, Jabalameli MS, Aryanezhad MB, Makui A (2013) Modeling and solving the bi-objective capacitated location-routing problem with probabilistic travel times. Int J Adv Manuf Technol 67:2007–2019

Ghiani G, Legato P, Musmanno R, Vocaturo F (2007) A combined procedure for discrete simulation-optimization problems based on the simulated annealing framework. Comput Optim Appl 38:133–145

Gibbons JD, Olkin I, Sobel M (1979) An introduction to ranking and selection. Am Stat 33(4):185–195

Golozari F, Jafari A, Amiri M (2013) Application of a hybrid simulated annealing-mutation operator to solve fuzzy capacitated location-routing problem. Int J Adv Manuf Technol 67:1791–1807

Grossmann IE (2014) Challenges in the application of mathematical programming in the enterprise-wide optimization of process industries. Theor Found Chem Eng 48(5):555–573

Hassan-Pour H, Mosadegh-Khah M, Tavakkoli-Moghaddam R (2009) Solving a multi-objective multi-depot stochastic location-routing problem by a hybrid simulated annealing algorithm. Proc Inst Mech Eng Part B: J Eng Manuf 223(8):1045–1054

Hassanzadeh A, Mohseninezhad L, Tirdad A, Dadgostari F, Zolfagharinia H (2009) Location-routing problem. In: Farahani R, Hekmatfar M (eds) Facility location: concepts, models, algorithms and case studies. Springer, Berlin, pp 395–417

Hemmati Golsefidi A, Akbari Jokar MR (2020) A robust optimization approach for the production-inventory-routing problem with simultaneous pickup and delivery. Comput Ind Eng 143:106388

Hong LJ, Fan W, Luo J (2021) Review on ranking and selection: a new perspective. Front Eng Manag 8(3):321–343

Klibi W, Lasalle F, Martel A, Ichoua S (2010) The stochastic multiperiod location transportation problem. Transp Sci 44(2):221–237

Kulcar T (1996) Optimizing solid waste collection in Brussels. Eur J Oper Res 90(1):71–77

Labbé M, Laporte G (1986) Maximizing user convenience and postal service efficiency in post box location. Belg J Oper Res Stat Comput Sci 26(2):21–35

Laganà D, Legato P, Pisacane O, Vocaturo F (2006) Solving simulation optimization problems on grid computing systems. Parallel Comput 32:688–700

Laporte G, Musmanno R, Vocaturo F (2010) An adaptive large neighbourhood search heuristic for the capacitated arc-routing problem with stochastic demands. Transp Sci 44(1):125–135

Legato P, Mazza RM, Vocaturo F (2021) Queueing, simulation and optimization for performance-oriented design of warehouse systems. In: Bruzzone AG, De Felice F, Massei M, Solis A (eds) Proceedings of the 20th international conference on modelling and applied simulation (MAS 2021), pp 141–151

Lopes RB, Ferreira C, Santos BS, Barreto S (2013) A taxonomical analysis, current methods and objectives on location-routing problems. Int Trans Oper Res 20(6):795–822

Mara STW, Kuo RJ, Asih AMS (2021) Location-routing problem: a classification of recent research. Int Trans Oper Res 28(6):2941–2983

Mirzaei-Khafri S, Bashiri M, Soltani R, Khalilzadeh M (2020) A robust optimization model for a location-arc routing problem with demand uncertainty. Int J Ind Eng Theory Appl Pract 27(2):288–307

Murty KG, Djang PA (1999) The U.S. army national guard's mobile training simulators location and routing problem. Oper Res 47(2):175–182

Nagy G, Salhi S (2007) Location-routing: issues, models and methods. Eur J Oper Res 177(2):649–672

Perl J, Daskin MS (1984) A warehouse location-routing methodology. J Bus Logist 5:92–111
Prodhon C, Prins C (2014) A survey of recent research on location-routing problems. Eur J Oper Res 238(1):1–17
Rand G (1976) Methodological choices in depot location studies. Oper Res Q 27(1):241–249
Renaud A, Absi N, Feillet D (2017) The stochastic close-enough arc routing problem. Networks 69(2):205–221
Sahinidis NV (2004) Optimization under uncertainty: State-of-the-art and opportunities. Comput Chem Eng 28(6–7):971–983
Salhi S, Rand GK (1989) The effect of ignoring routes when locating depots. Eur J Oper Res 39(2):150–156
Sangaiah AK, Tirkolaee EB, Goli A, Dehnavi-Arani S (2020) Robust optimization and mixed-integer linear programming model for LNG supply chain planning problem. Soft Comput 24:7885–7905
Schmidt G, Wilhelm WE (2000) Strategic, tactical and operational decisions in multi-national logistics networks: a review and discussion of modelling issues. Int J Prod Res 38(7):1501–1523
Schneider M, Drexl M (2017) A survey of the standard location-routing problem. Ann Oper Res 259(3):389–414
von Boventer E (1961) The relationship between transportation costs and location rent in transportation problems. J Reg Sci 3(2):27–40
von Lanzenauer CH, Pilz-Glombik K (2002) Coordinating supply chain decisions: An optimization model. OR Spectrum 24:59–78
Wang Y, Peng S, Zhou X, Mahmoudi M, Zhen L (2020) Green logistics location-routing problem with eco-packages. Transp Res Part E: Logist Transp Rev 143:102118
Wasner M, Zäpfel G (2004) An integrated multi-depot hub-location vehicle routing model for network planning of parcel service. Int J Prod Econ 90(3):403–419
Watson-Gandy C, Dohrn P (1973) Depot location with van salesmen—a practical approach. Omega 1(3):321–329
Wilcox R (1984) A table for Rinott's selection procedure. J Qual Technol 16(2):97–100
Yoon M, Bekker J (2019) Considering sample means in Rinott's procedure with a Bayesian approach. Eur J Oper Res 273(1):249–258
Zarandi MHF, Hemmati A, Davari S (2011) The multi-depot capacitated location routing problem with fuzzy travel times. Expert Syst Appl 38(8):10075–10084
Zhang S, Chen M, Zhang W, Zhuang X (2020) Fuzzy optimization model for electric vehicle routing problem with time windows and recharging stations. Expert Syst Appl 145:113123