The life and times of the Savings Method for Vehicle Routing Problems

Graham K Rand

Received: 1 September 2009; Revised: 6 October 2009; Accepted: 18 October 2009

Abstract
Forty five years ago, an academic and practitioner from the north of England published a method of tackling the vehicle routing problem (VRP) in an American journal. Little could they have realised how the method they devised would still be a significant part of the research agenda nearly half a century later. Adaptations of their method are significant components in the analysis of the many different extensions to the problem that have been investigated. This paper provides the historical background to the development of the savings method and subsequent proposed variations to the basic savings formula and other improvements, and then charts the role the savings method has played in the investigation of VRPs with additional constraints. Some interesting examples of practical applications of the savings method are reported. Finally, comments are made on the use of the savings method in commercial routing packages.

Key words: Savings method, Clarke and Wright, vehicle routing, logistics.

1 Introduction

The vehicle routing problem (VRP) can be stated very simply: a set of customers with known location and demand are to be supplied from a depot by delivery vehicles of known capacity subject to all customer demand being met, vehicle capacity not being exceeded and total trip length not exceeding some specified level. If times of delivery at customers’ locations are also specified then it is often described as the “vehicle scheduling” problem. The routes begin and end at the depot. A customer’s demand is satisfied by a single vehicle in a single delivery. The constraints arise because (i) it is assumed all demand must be met that day, although this is not always the case, (ii) there is a limit on how much product the vehicle can carry, but it is very difficult to define this precisely, and (iii) there may be a legal restriction on the length of the trip in hours. The VRP has been shown to be NP-hard (Lenstra and Rinnooy Kan, 1981) and therefore it is unlikely that a polynomial-time algorithm will be found for this problem. An amusing, but salutary, story of why this may not be a sufficient description of the VRP, in the context of applying the savings method, is provided by Woolsey (1991).
The vehicle routing problem was first described by Dantzig and Ramser (1959), who provided a solution method based, not surprisingly, since one of the authors became known as the “father of linear programming” (Gass, 2003), on linear programming. Clarke and Wright (1964) developed a heuristic solution method which became known as the savings method and was the first algorithm that became widely used. It is also called the Clarke-Wright algorithm, after the authors, but in the early years it was also described as the Wright-Fletcher-Clarke algorithm (Unwin, 1968) or the Fletcher-Clarke-Wright algorithm (Knight and Hofer, 1968; Yellow, 1970). Fletcher and Clarke had given a paper at the 1963 conference of the Operational Research Society held in Nottingham (Fletcher and Clarke, 1963). Surprisingly, their heuristic obtained a better result than Dantzig and Ramser for the illustrative example that they had used (Clarke and Wright, 1964). Dantzig and Ramser’s approach had been to consider linking pairs of customers into a route that were close together, just considering the distance between them, but Clarke and Wright extended this, as shall be seen, to take into account the reduction in distance obtained by linking two customers into a route, rather than serving them on separate routes.

Clarke was employed by the Cooperative Wholesale Society in Manchester, England (as was Fletcher), though by the time the paper was published he had moved to work at ICI in Hyde. Wright was at the University of Manchester, then the Manchester College of Science and Technology. It was he who had pointed out the paper by Dantzig and Ramser (1959) to Clarke and Fletcher. Fletcher and Clarke (1963) give credit to Wright for developing “almost the same method as the one described in this paper.” The original implementations were in FORTRAN for an IBM 7090 and in AUTOMATH for a Honeywell 800 (Fletcher and Clarke, 1963). The specific example given in the paper by Clarke and Wright (1964) involved 30 customers being served from a depot in Newton Heath, a suburb of Manchester. For a single instance, the current practice resulted in 10 routes with a total mileage of 1 766, which the savings method improved to 8 routes and 1 427 miles.

The Clarke-Wright paper concludes with the tantalising comment that “details of some of these restrictions together with computational methods for a digital computer will be found in a Case Study which will be published shortly.” There is no evidence of such a paper being published.

2 The savings method

The savings method first calculates “savings” for every pair of customers. If there are two customers $i$ and $j$, respectively at a distance $c_{i0}$ and $c_{0j}$ from the depot, and distant $c_{ij}$ from each other, then if the deliveries are made separately the total distance is, when distances are symmetric, $2c_{i0} + 2c_{0j}$, but if the deliveries are combined into a route, then the total distance is $c_{i0} + c_{0j} + c_{ij}$; hence there is a “saving” in mileage of $s_{ij} = c_{i0} + c_{0j} − c_{ij}$.

The savings thus calculated are then ranked and customers are placed on routes into which they can be linked (otherwise a new route is started) until a constraint (maximum route length, vehicle capacity) is reached.

Consider the following example. Ten customers are to be served from a depot at location (150, 250) by vehicles each with a capacity of 3.5 tonnes. The customer details are given in Table 1.
The life and times of the Savings Method for Vehicle Routing Problems

For convenience, the distances between the depot and customers, and between customers, are calculated using Pythagoras' Theorem and given in Table 2. The savings are then calculated as indicated previously, and presented in Table 3. The assumption of the method is that the higher the savings, the better it is to link those two customers on a route. For instance, customers 3 and 4 are far from the depot and close together. It makes sense to link those customers into one route. The savings are ranked and a decision is made at each point as to whether the link can be placed into a route. Links can only be made where no constraint is broken by the route created by making that link. Furthermore, links can only be made if proper routes are formed, so when amending routes already created, new links can only be added to the start or end of routes.

<table>
<thead>
<tr>
<th>Locations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>23.26</td>
<td>29.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>25.94</td>
<td>32.28</td>
<td>2.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>20.81</td>
<td>14.56</td>
<td>33.73</td>
<td>35.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>18.68</td>
<td>19.45</td>
<td>17.89</td>
<td>18.97</td>
<td>17.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>24.17</td>
<td>19.10</td>
<td>33.24</td>
<td>34.48</td>
<td>5.39</td>
<td>15.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>26.57</td>
<td>27.80</td>
<td>19.21</td>
<td>19.10</td>
<td>22.47</td>
<td>8.06</td>
<td>19.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>30.46</td>
<td>25.32</td>
<td>37.59</td>
<td>38.48</td>
<td>11.00</td>
<td>19.72</td>
<td>6.32</td>
<td>21.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>32.14</td>
<td>31.02</td>
<td>28.43</td>
<td>28.28</td>
<td>21.21</td>
<td>14.14</td>
<td>16.4</td>
<td>9.22</td>
<td>15.52</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Distance matrix for illustrative example.

<table>
<thead>
<tr>
<th>Locations</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>19.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>11.18</td>
<td>4.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10.19</td>
<td>4.02</td>
<td>39.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.57</td>
<td>12.99</td>
<td>0.01</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.02</td>
<td>0.02</td>
<td>8.17</td>
<td>8.77</td>
<td>1.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>3.48</td>
<td>8.71</td>
<td>0.76</td>
<td>1.21</td>
<td>21.56</td>
<td>3.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.07</td>
<td>0.01</td>
<td>14.79</td>
<td>16.59</td>
<td>4.47</td>
<td>11.20</td>
<td>7.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2.68</td>
<td>8.00</td>
<td>1.92</td>
<td>2.71</td>
<td>21.45</td>
<td>5.04</td>
<td>26.38</td>
<td>11.49</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.01</td>
<td>1.30</td>
<td>10.09</td>
<td>11.92</td>
<td>10.24</td>
<td>9.63</td>
<td>15.31</td>
<td>22.49</td>
<td>21.69</td>
</tr>
</tbody>
</table>

Table 3: Savings matrix for illustrative example.

The approach adopted here is a “parallel” one. When a link cannot be made into an existing route a new route is immediately created with that link. In this illustrative example
the only constraint being considered is the capacity of the vehicles. The application of the method is presented in Table 4.

<table>
<thead>
<tr>
<th>Saving</th>
<th>Customers</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>39.66</td>
<td>3 4</td>
<td>Total demand 2.2 — OK</td>
</tr>
<tr>
<td>26.38</td>
<td>7 9</td>
<td>Total demand 1.3 — OK</td>
</tr>
<tr>
<td>22.49</td>
<td>8 10</td>
<td>Total demand 0.7 — OK</td>
</tr>
<tr>
<td>21.69</td>
<td>9 10</td>
<td>Combines the routes with total demand 2.0 — OK</td>
</tr>
<tr>
<td>21.56</td>
<td>5 7</td>
<td>Breaks constraint</td>
</tr>
<tr>
<td>21.45</td>
<td>5 9</td>
<td>Breaks constraint</td>
</tr>
<tr>
<td>19.70</td>
<td>1 2</td>
<td>Total demand 1.8 — OK</td>
</tr>
<tr>
<td>16.59</td>
<td>4 8</td>
<td>Breaks constraint</td>
</tr>
<tr>
<td>15.31</td>
<td>7 10</td>
<td>Already in route</td>
</tr>
<tr>
<td>14.79</td>
<td>3 8</td>
<td>Breaks constraint</td>
</tr>
<tr>
<td>12.99</td>
<td>2 5</td>
<td>Breaks constraint</td>
</tr>
<tr>
<td>11.92</td>
<td>4 10</td>
<td>Breaks constraint</td>
</tr>
<tr>
<td>11.49</td>
<td>8 9</td>
<td>Already linked</td>
</tr>
<tr>
<td>11.20</td>
<td>6 8</td>
<td>Link customer 6 into second route, total demand 2.4 — OK</td>
</tr>
</tbody>
</table>

Table 4: Steps in the application of the savings method for the illustrative example.

The method would continue to proceed through all the savings, but here we can see that customer 5 has to be on a separate route, because the demand would break the capacity constraint on all the routes already established. The result is given in Table 5 and illustrated in Figure 1.

<table>
<thead>
<tr>
<th>Route</th>
<th>Load</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depot – 4 – 3 – Depot</td>
<td>2.2</td>
<td>45.32</td>
</tr>
<tr>
<td>Depot – 7 – 9 – 10 – 8 – 6 – Depot</td>
<td>2.4</td>
<td>58.38</td>
</tr>
<tr>
<td>Depot – 1 – 2 – Depot</td>
<td>1.8</td>
<td>36.79</td>
</tr>
<tr>
<td>Depot – 5 – Depot</td>
<td>2.1</td>
<td>26.68</td>
</tr>
<tr>
<td>8.5</td>
<td>167.17</td>
<td></td>
</tr>
</tbody>
</table>

Table 5: The result from the application of the savings method for the illustrative example.

An alternative to the “parallel” approach is a “sequential” one, in which one route at a time is developed. Results (Laporte and Semet, 2002) indicate that the parallel version is clearly better than the sequential approach. However, as this is a heuristic there is no guarantee that optimal, or even near optimal, results will be obtained. In this example, for instance, it can be seen, by inspection, that the three routes given in Table 6 are better as one fewer vehicle is required and the travel distance is reduced.

<table>
<thead>
<tr>
<th>Route</th>
<th>Load</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depot – 3 – 4 – 8 – 10 – 6 – Depot</td>
<td>3.3</td>
<td>71.35</td>
</tr>
<tr>
<td>Depot – 5 – 7 – 9 – Depot</td>
<td>3.4</td>
<td>44.15</td>
</tr>
<tr>
<td>Depot – 1 – 2 – Depot</td>
<td>1.8</td>
<td>36.79</td>
</tr>
<tr>
<td>8.5</td>
<td>152.29</td>
<td></td>
</tr>
</tbody>
</table>

Table 6: Better results for the illustrative example.

This effect arises from a myopic feature of the algorithm: once a link is selected it cannot
be deselected. Holmes and Parker (1976) developed an algorithm designed to overcome this weakness by allowing the temporary prohibition of links that yield high savings, but which affect subsequent links. Whilst this gave encouraging results on their test problems, it would appear that the approach was not attractive to subsequent researchers, perhaps because of the computational implications.

It has been argued (Cordeau et al., 2002b) that four attributes of good VRP heuristics are accuracy, speed, simplicity and flexibility. The savings method scores highly on speed and simplicity, as it contains no parameters and is easy to code. Extensive testing of algorithms show that it is not the most accurate, measured as a percentage of the optimal value, and the authors assess it as being average on this attribute. Their assessment of its flexibility, the ability to incorporate side constraints encountered in real-life applications, is that it is the worst feature. Additional constraints can be incorporated, but a significant deterioration in accuracy results.

3 Modifications to the savings method

Several authors proposed developments of the algorithm. These developments may be categorised as adaptations to the savings formula, methods to speed up computation time and improvements to the route merging process.

Gaskell (1967) was concerned about the tendency of the savings method to produce peripheral routes, particularly when the capacity constraint is important. He proposed that the between-customer distance should be adjusted by a route shape parameter $\lambda$ when calculating savings, so that the savings formula becomes $s_{ij} = c_{i0} + c_{0j} - \lambda c_{ij}$. As the parameter $\lambda$ increases from zero, more emphasis is placed on the distance between the customers rather than their distances to the depot. However, the question arises as to
what value of $\lambda$ should be used in a practical study. McDonald (1972) showed that with any fixed $\lambda$, results which are far from optimum may be obtained, and that there was no value of $\lambda$ which was significantly better than any other value.

Paessens (1988) proposed a savings function of the form $s_{ij} = c_{i0} + c_{0j} - \lambda c_{ij} + \mu |c_{0i} - c_{j0}|$ with the parameters $\lambda$ and $\mu$ being in the ranges $0 < \lambda \leq 3$ and $0 \leq \mu \leq 1$, respectively. No motivation is provided for this modification. Altinel and Öncan (2005) observed that the merging of routes with equal or nearly equal savings occurs frequently, particularly towards the end of the process. They therefore proposed a variation which takes into account the customer demands when calculating savings. Their saving criterion becomes $s_{ij} = c_{i0} + c_{0j} - \lambda c_{ij} + \mu |c_{0i} - c_{j0}| + \nu (d_i + d_j)/d$ where $d_i$ is the demand of customer $i$, $\bar{d}$ is the average customer demand, and $\nu$ is a new parameter. This requires determining the best values of three parameters, and therefore, in their computational experiments, they evaluated nearly 9,000 different parameter vectors and obtained a considerable improvement compared to the original savings method. Subsequently, Battara et al. (2008) developed a genetic algorithm to determine the best sets of parameters, which gave results of similar quality with much shorter computing times.

Yellow’s (1970) focus was on speeding up computation time, and he argued that the need to determine a savings matrix, and then rank the savings, could be removed by searching a list of the customers ordered by polar coordinates. Golden et al. (1977) were concerned about the need to store distances when varying the route shape parameter, and proposed that only certain linkages should be selected, rather than all the links between pairs of customers. They also proposed a sorting procedure using a heap data structure.

Nelson et al. (1985) contributed to speeding up the algorithm by using special data structures to make the determination of the maximum saving value computationally more efficient. Six methods for implementing the savings method were presented and compared: (1) a standard method using a shellsort; (2) using a heapsort; (3) and (4) alternative ways of explicitly storing pointers to eliminate arcs that are associated with nodes that can no longer be considered; (5) assuming a complete graph and storing and solving the same problems as (3) and (4) with less storage and computational time; and (6) using a number of sequentially created heaps to store the savings. In (6) savings above a threshold value are stored in a heap and processed to completion using the data structures of (4). Subsequent heaps are created only for nodes that have not become interior to emerging routes. The authors concluded that use of heap structures reduces computational time, as does the use of pointers in methods (3) and (4). Method (5) realised a substantial reduction in storage, without loss of computational efficiency, and (6) produced the shortest run times in all cases and smallest storage requirements in most cases. However, they recommended caution in applying (6) in practice because of the artificiality of the test problems.

Paessens (1988) proposed methods for storing distances in a 1-dimensional array and determination of the maximum saving value. Three sorting strategies were considered, but the results on a number of test problems indicate that it is difficult to recommend a particular approach.

Other authors have sought to optimise the route merging process. Altinkemer and Gavish (1991) replaced the merging procedure of the savings method by a matching procedure which merges partial solutions at each step. At each iteration of the algorithm, multiple
clusters of nodes are merged, the number of clusters being determined by solving a matching problem, which maximises the savings obtained. Their algorithm is polynomial, with a time complexity of $O(n^3)$, where $n$ denotes the number of customers. Wark and Holt (1994) proposed a method in which routes were also formed by growing clusters of customers, but which considers two different ways of merging pairs of clusters. The process starts with each customer in its own cluster, before new cluster sets are formed by one of two matching rules, based on either the “cost” of the cluster, which is either a function of the route associated with the cluster alone, or a function of load and timespan.

A recent development has been the application of Ant Systems to the VRP, using modifications of the savings method (Reimann et al., 2004). In such systems, a population of artificial agents repeatedly constructs solutions to the problem using a joint population memory and some heuristic information. After each member of the population has constructed its next solution, the memory is updated with a bias towards better solutions found. Gradually, the memory will build up, thus giving stronger influence to the solutions built by the artificial agents, and the solutions will evolve towards the global optimum. Reimann et al. (2004) modified the savings method to create a savings-based ant system that not only improves “the efficiency, but also improves the effectiveness of the algorithm leading to a fast and powerful problem solving tool for real world sized Vehicle Routing Problems.”

4 The savings method applied to variations of the VRP

Many variations of the VRP have been addressed over the years. An obvious variation, if indeed it is such, is one in which it is acknowledged that the road network is not symmetric: the Asymmetric Vehicle Routing Problem. This applies particularly to urban areas, where distances may be affected by one-way systems. Vigo (1996) addresses this problem, together with the standard constraint of vehicle capacity, and acknowledges that “the extension of the Clarke-Wright algorithm to asymmetric instances poses no difficulty.” As the distance matrix is asymmetric, so also will be the savings matrix. Therefore, a link $(i,j)$ can be considered for joining two routes only if customer $i$ (or $j$) is the last customer of a developed route and vertex $j$ (or $i$) is the first of the other, thus halving the joining possibilities when compared to the symmetric situation.

Another obvious extension to the basic formulation of the VRP is to consider more than one depot. This problem was investigated originally by Tillman and Cain (1972), following an earlier paper (Tillman, 1969), and by Wren and Holliday (1972). The application of the savings method in this situation is not straightforward. Consider determining the savings achieved when linking two demand points which are close to one depot, so naturally falling within that depot’s “area,” and some distance from a second depot. The savings when calculated from the second depot will be greater than the savings from the first, more obvious, depot from which to serve these customers. Wren and Holliday developed an algorithm that did not depend on the savings method, because of concerns about computer time required, but Tillman and Cain adapted the savings formula, by reducing the savings when calculated from the further depot by the amount that the actual distance exceeds the distance to the nearer depot. To achieve this, when there are $m$ depots, a
modified distance is calculated from each depot, \( k \), to each customer, \( i \), by the expression 
\[
  c_{ik}^{\text{mod}} = \min_m \{c_{im}\} - (c_{ik} - \min_m \{c_{im}\})
\]
before the savings from each depot is determined by the usual formula using the modified distances. The branch-and-bound algorithm proposed by Tillman and Cain uses an upper bound based on the maximum saving.

An interesting variation of the VRP is presented by Basnet \textit{et al.} (1999), who consider the VRP with tree-like road networks that occur, for example, in rural road systems where customers are located on rural roads leading off from a few highways which form the “trunks” of a tree-like network. Of the heuristics developed, the one that provided the best solutions was a modification of the savings method.

However, there are several well-established variations of the VRP, and it is these that will be examined in the following sections, with a view to understanding how effectively the savings method contributes to the search for good solutions.

4.1 With time windows

When the delivery at some, or all, customers must be within a pre-specified time window, the VRP becomes the \textit{VRP with Time Windows} (VRPTW). Soft time windows can be broken, but at a cost, while hard time windows do not allow a vehicle to arrive outside the window. Most published work has investigated the problem with hard time windows. Surveys of this problem are provided by Desrochers \textit{et al.} (1988), Cordeau \textit{et al.} (2002a) and Bräysy and Gendreau (2005a, 2005b). Solomon (1987) observes that the early work on the VRPTW was grounded in practical studies.

The savings method has proved a useful building block in the construction of algorithms for this problem. Two issues arise. It is necessary to take into account the timing of a proposed link, and the route direction, which is not relevant in a symmetric distance VRP without time window constraints. Two partial routes with end customers \( i \) and \( j \), respectively, have compatible orientations if \( i \) is either the first or last customer and \( j \) is either the last or first customer, so that the links are from the last customer on one route to the first customer on the other. In addition to taking into account capacity constraints, the time window constraints must be checked at each step in the process. It may appear to be advantageous, when determining savings, to join two customers close in distance but far apart in time, giving a long period of waiting time, which is potentially costly as the vehicle could be serving other customers rather than waiting.

Solomon’s (1987) initial approach to solving the VRPTW was to extend the savings method, using the parallel savings method with the list processing and heapsort structures proposed by Golden \textit{et al.} (1977). To overcome the waiting time problem, Solomon suggests introducing a constraint to limit the waiting time. Homberger and Gehring (1999) developed two “evolutionary” strategies for the VRPTW. Both methods require a starting solution, which is generated by a stochastic approach based on the savings method.

A bicriteria heuristic based on the savings method, in which the original savings measure in terms of travel time is combined with “loss of flexibility,” was developed by van Landeghem (1988). Although deliveries outside the time window are not allowed, a vehicle can arrive before the opening time and wait. After the delivery has been completed, it must leave immediately for its next destination. Thus waiting at a customer site occurs only before
the opening time of the next time window. An upper limit can be put on this waiting time, as well as on the total waiting time over the complete route. The larger this upper limit, the more flexibility will be present in the route building process, resulting in shorter routes, but possibly taking more elapsed time.

Balakrishnan (1993) described three heuristics for the problem with soft time windows, in which the time window constraints may be broken, but at a cost. One heuristic was based on the savings method. By allowing limited time window violations for some customers, it may be possible to obtain significant reductions in the number of vehicles required and/or the total distance or time of all routes.

Atkinson (1994, 1998) published two papers on the VRPTW, motivated by some previous work concerning the delivery of school meals (Atkinson, 1990). The first (1994) described a greedy heuristic for a class of combinatorial optimisation problems, which was demonstrated by application to the VRPTW. A key component of the method was a look-ahead capability, which gave a performance improvement that was at least as great as, and in addition to, that which had been obtained through use of the savings method. In the second paper (1998), a greedy randomised heuristic, based on the 1994 heuristic and incorporating a version of the savings method, was developed.

Bräysy (2002) describes several local search heuristics for the VRPTW. One of the heuristics, a so-called merge heuristic, draws its basic concepts from the savings method, in which the original measure of savings is modified to allow for changes in waiting times. Moreover, the customers in the combined route are reordered before evaluating the savings incurred by uniting the two routes.

4.2 With backhauls

When the route of the vehicle contains deliveries (linehauls) and then collections (backhauls) the VRP becomes the VRP with Backhauls (VRPB). Surveys of this problem are provided by Casco et al. (1988) and Toth and Vigo (2002). The VRPB has become increasingly important because integrating backhauls with linehauls utilises the vehicle capacity and reduces cost.

The first approaches to this problem extended the savings method. Deif and Bodin (1984) first introduced a constraint to ensure that all deliveries are made before any collection. Then, in a second extension, they delayed backhaul customers from early inclusion in routes by introducing a penalty factor in the basic savings formula. They experimentally demonstrated that the best results are obtained when the formation of mixed routes is delayed. So the saving definition was amended to penalize the arcs connecting service points of the two types of customers, thus delaying the union of linehaul and backhaul routes. The backhaul saving is defined as $S_{ij} = s_{ij} - pM$ where $i \in L; j \in B$ ($L$ and $B$ are respectively the linehaul and backhaul customer subsets) or \textit{vice-versa}, where $M$ is an estimate of the maximum saving $s_{ij}$, and $p$ is a real penalty multiplier between 0 and 1, or otherwise it is $s_{ij}$. The drawback of these two approaches was that the number of routes could not be controlled in the final solution. Hence, for a given instance the solution found may require more vehicles than the maximum available to serve all the customers.

Other extensions of the savings method for this problem have been presented by Golden
et al. (1985) and Casco et al. (1988); the former treats the case in which the precedence constraints between linehaul and backhaul vertices are not present.

Wassan (2007) presented an algorithm in which the initial solution is generated using two construction schemes that had been developed by Osman and Wassan (2002), based on savings, insertion and assignment approaches, using the savings concept, with some modifications to generate the required number of vehicle routes. The first of these schemes was called the saving-insertion heuristic (SIH), which starts by building routes containing only customers receiving a delivery, before the backhauls are then inserted into the generated routes according to a best-cost insertion criterion subject to the customer and vehicle capacity constraints. The second scheme, called the saving-assignment heuristic (SAH) starts by generating feasible vehicle routes for both sets of delivery and backhaul customers separately, by using the SIH procedure, followed by an improvement process, before both routes are then merged.

4.3 With pickups

The VRPB requires that the delivery (linehaul) and pickup (backhaul) customers are different, but demands for packaging and used product returns from customer locations have increased substantially due to environmental and government regulations, so that a customer may have both delivery and pickup demands. This gives rise to vehicle routing problems with simultaneous pickups and deliveries (VRPPD). An assumption here is that a customer would prefer to have a single stop rather than multiple stops of a vehicle for convenience, and that is likely to be desirable from the perspective of the transport company, even though relaxing the constraint might improve the solution to the problem. A survey of the VRPPD was conducted by Desaulniers et al. (2002). If people are being transported rather than goods, then this is the so-called dial-a-ride problem, which motivated the early work on VRPPD.

Gronalt et al. (2003) address a version of this problem, using heuristics based on the savings method. In their problem, customers place orders requiring shipments between two locations with the transportation company. The transportation company serves these orders from a number of distribution centres. So goods are shipped between the pickup location of an order and the closest distribution centre, between distribution centres, and between a distribution centre and the delivery location of an order. Loads between distribution centres and pickup or delivery locations are generally less-than-truckload, but movements between different distribution centres are concerned with full truckloads, because of load consolidation. Furthermore, within the planning horizon each vehicle can be assigned more than one tour.

The authors present four algorithms. The first is a straightforward adaptation of the savings method, which cannot be used in its basic form because there are multiple depot pickups and deliveries. So an algorithm is developed, in which savings are calculated between orders, rather than between locations, because two locations are associated with each order. The second algorithm is called the opportunity savings algorithm, which, together with the third algorithm, takes into account some measure of opportunity costs associated with the choice of an order combination. It is argued that the attractiveness of joining two orders may depend on the alternative options these orders have as well as
the savings value associated with these two orders. So, after the calculation of all the savings, the best and the second best choice for each order is found, and the difference in the savings is interpreted as opportunity costs of not choosing the best option. A new measure is determined by adding some fraction of these opportunity costs to the savings value associated with the best choice. The third algorithm, called the simultaneous-savings algorithm, looks at the structure of the savings matrix in each iteration and anticipates that this structure will change from iteration to iteration. The final algorithm combines this simultaneous approach with the opportunity cost approach.

The authors’ results showed that these heuristics find very good solutions quickly. In particular, they found that using a measure for opportunity costs significantly improves the solution quality, while a simultaneous approach does not have a favourable effect.

4.4 With stochastic demands

In the VRP with stochastic demands (VRPSD), the customers’ demands are stochastic variables, for which only the probability distribution for each customer is assumed known at the time of planning, so that it is the expected total travel cost which is subject to minimization. The total actual demand on a route may exceed the vehicle capacity, in which case a failure is said to occur. A strategy is required for updating the routes when this happens, and the action resulting from this strategy is called a recourse action. As the strategy adopted affects the expected cost of a given route, it must be known at the time of planning. Reviews of work on the VRPSD are provided by Dror and Trudeau (1986), Dror et al. (1989) and Gendreau et al. (1996).

The first to address the VRPSD was Tillman (1969) who based his solution approach on a modification of the savings method in the context of a multi-depot variant of the VRP with Poisson distributed demands. The model considered a cost trade-off between exceeding the vehicle capacity and completing the route with excess capacity.

Stewart and Golden (1983) presented three models for the VRPSD, and considered several demand distributions. They compared two heuristics, one of which is based on the savings method. They stated that the savings method is easily adapted to this problem and runs very quickly. The VRP capacity constraints are replaced with other constraints to decide whether two current tours may be feasibly joined. If the probability of route failure is too high, this combination is disallowed. They also adapted the savings method to handle penalty function models. When a fixed penalty for route failure is modelled the savings function becomes

\[ s_{ij} = c_{0i} + c_{0j} - c_{ij} + \lambda P_i + \lambda P_j - \lambda P_{ij} \]

where \( s_{ij} \) is the savings realised by joining customers \( i \) and \( j \) on the same route, \( P_i \) is the probability that the route including customer \( i \) fails, \( P_{ij} \) is the probability that the combined route with both \( i \) and \( j \) fails, and \( \lambda \) is the penalty incurred when a route fails. When a cost is incurred for each unit of demand left unsatisfied, the savings function is defined as

\[ s_{ij} = c_{0i} + c_{0j} - c_{ij} + \lambda \ell_i + \lambda \ell_j - \lambda \ell_{ij} \]

where \( \ell_i \) and \( \ell_{ij} \) are the expected number of units short on the routes including customer \( i \) and the route \( i \) and \( j \) will be on, respectively, and \( \lambda \) is the penalty per unit short. Both \( P_{ij} \) and \( \ell_{ij} \) may be calculated knowing the probability distributions of demand. These quantities must be updated every time two customers are joined. In both variants of the penalty function model, customers are linked strictly according to savings; the capacity constraints are ignored.
4.5 With stochastic travel times

Laporte et al. (1992) introduced the VRP with stochastic travel and service times. Lambert et al. (1993) adapted the savings method for this problem, in the context of optimizing money collection routes through bank branches in a network with stochastic travel times. These branches were located in Belgium, with headquarters at Namur. In addition to the standard VRP, there was an unusual cost aspect: money received at a branch between the vehicle’s departure and closing time loses a day’s interest as it is not collected until the next day. Thus visits to branches producing a large amount of cash should be made as late in the day as possible. The objective function takes into account penalties associated with lost interest. The savings method was extended by considering more general objective functions that do not simply minimize routing costs, but include other terms as well (such as penalties for lost interest). Consider two vehicle routes $l_1$ and $l_2$ including arcs $(i, 0)$ and $(0, j)$ respectively, and the costs $z(l_1)$ and $z(l_2)$ of these routes. Merging $l_1$ and $l_2$, by introducing $(i, j)$ and deleting $(i, 0)$ and $(0, j)$ yields a route $l$ of cost $z(l)$. If $l$ is feasible and the saving $s_{ij} = z(l_1) + z(l_2) - z(l)$ is positive, the merge is implemented. The authors present results for networks with 28 and 44 nodes. An aspect of this work that aroused the interest of the bank’s management was that major branches tended to be located towards the end of the routes, as the objective function contains a term for lost interest. Such solutions do not arise naturally if standard VRP algorithms are used.

5 The savings method and the fleet mix problem

The composition of a vehicle fleet is an important cost investment decision, as once a vehicle, or fleet of vehicles, is purchased, it will be kept for some time, even if it proves to have been an unwise decision. A logistics manager may need to determine the best composition of the fleet in order to respond to customer orders at minimal cost. The appropriate fleet need not be homogeneous and a good vehicle fleet mix is likely to yield better results. Researchers have sought to bring some insight to this problem and develop an efficient heuristic, which incorporates both routing and vehicle fixed costs, to produce a number of good feasible fleet mixes.

Golden et al. (1984) were the first to address this problem. They developed several algorithms, some of which are based on the savings method. They extended the concept of savings to include fixed vehicle costs. The total cost of a subtour is the sum of its routing costs and the cost of the vehicle used. When a larger subtour is created from two subtours, the total cost of this larger subtour can be derived in the same way. To choose the two subtours which should be merged, the subtours which generate the greatest total savings when merged need to be determined. The combined savings when joining customers $i$ and $j$ ($i$ and $j$ are one of the endpoints of subtours $I$ and $J$ respectively) is $S_{ij} = c_{i0} + c_{0j} - c_{ij} + F(z_i) + F(z_j) - F(z_i + z_j)$ where $F(z)$ is the fixed cost of the smallest vehicle that can service a demand of size $z$. The total demands on the two subtours are $z_i$ and $z_j$. These savings must be calculated at every step. This method (denoted as Combined Savings (CS)) is the basis for the remaining savings algorithms developed by the authors.

The fact that $S_{ij}$ represents only the immediate savings gained by joining customers $i$ and $j$
means that the procedure is somewhat imprecise. Therefore the authors developed variants of the method, called opportunity savings algorithms, which include opportunity savings in the calculation of total savings. The opportunity savings are a function of the unused capacity of the vehicle serving the merged subtours. In the first of these variants, called Optimistic Opportunity Savings (OOS), the opportunity savings are defined to be the cost of the smallest vehicle that can serve all the unused capacity of the new vehicle. A second variant, called Realistic Opportunity Savings (ROS), results from two considerations. First, opportunity savings should not be included in the savings formula unless the merging of two subtours requires the use of a larger vehicle as the purpose of opportunity savings should be to encourage the use of larger vehicles when it seems profitable to do so. Therefore the authors introduce a vehicle threshold, which is crossed if the vehicle which would service the merged subtours is larger than each of the vehicles presently used. In the second variant, opportunity savings are only included if the threshold is crossed. The second consideration arises from their computational experiments, in which the total demand of a subtour tended to be close to a vehicle capacity. Therefore, they argued that it is probably best to use the largest suitable vehicle, which is the vehicle just smaller than that given by the OOS variant, except when the unused capacity equals a vehicle capacity, in which case both algorithms use the same vehicle.

Computational results showed that none of the above methods are powerful enough that a single application will consistently generate good solutions. It is therefore necessary to vary these algorithms so that they produce a number of different solutions for each problem, with the least costly of these being selected as a final solution. A method used which introduces variety into the ROS algorithm is referred to as ROS-γ, where γ is the route shape parameter, λ, of Gaskell (1967). ROS-γ yielded the best results among all the savings algorithms that were tested.

Desrochers and Verhoog (1991) developed a matching-based savings algorithm for this problem. The savings method is adapted: a new form of saving is proposed and the way the best combination of the subtours are merged is different. Their basic saving, denoted by DVS, is given by $DVS_{ij} = TSP(N_i) + TSP(N_j) - TSP(N_i U N_j)$, where $N_i$ denotes the set of customers in the $i$-th route and $TSP(N_i)$ the cost of the optimal travelling salesman tour in that route. For the choice of the best feasible combination of subtours to be merged, the maximum weighted matching problem, with the weights equal to the savings obtained above, is solved and the combination with the largest saving is chosen. This provides a less myopic criterion than the usual savings method.

Salhi and Rand (1993) investigated this problem, using a routing procedure based on the savings method (Salhi and Rand, 1987). This generates good routes for a given vehicle capacity and then several modules are introduced in turn to derive, whenever possible, more economical vehicle fleet mixes (those with a cheaper total cost) without violating constraints such as time restriction and capacity.

Liu and Shen (1999) and Dullaert et al. (2002) addressed an extended version of the problem, by considering time window constraints. Liu and Shen (1999) developed several insertion-based parallel savings heuristics capable of generating feasible solutions. Instead of merging individual routes, the insertion of each route, in its original or reversed order, is evaluated in all possible insertion places in all other routes for different parameter
settings. They modified Golden et al.'s (1984) savings criteria to take possible savings in vehicle acquisition costs into account. Their parallel savings heuristics are motivated by Solomon's (1987) sequential insertion heuristics. Instead of merging routes, one route is inserted into another. Dullaert et al. (2002) extended Solomon's (1987) sequential insertion heuristic II with vehicle insertion savings, based on Golden et al. (1984), and obtained significantly better solutions. Golden et al.'s (1984) savings formulations were adapted by considering the load of a vehicle and its maximum capacity. The authors define the Adapted Combined Savings (ACS) as the difference between the fixed costs of the vehicle capable of transporting the load of the route after and before inserting a customer. Their Adapted Optimistic Opportunity Savings extends the ACS by subtracting the fixed cost of the smallest vehicle that can serve the unused capacity. Finally, their Adapted Realistic Opportunity Savings takes the fixed cost of the largest vehicle smaller than or equal to the unused capacity into account as opportunity saving. It only does so if a larger vehicle is required to serve the current tour after a new customer has been inserted.

6 Application of the savings method and its variations

The motivation for the original paper was to improve the distribution of deliveries of Cooperative Wholesale Society vehicles in the English Midlands: 400 customers were served from Manchester (Fletcher and Clarke, 1963). Since then, many applications of the savings method to real problems have been reported. The following descriptions draw attention to a selection of the more interesting of them.

McDonald (1972) describes the application of Gaskell’s modified savings method (Gaskell, 1967) to the problem of collecting medical specimens. To overcome the method’s weakness, that once a link is made it is never removed, the 3-optimal method (Lin, 1965) was used to determine whether any improvement could be made to the best results obtained by Gaskell’s method. An immediate result of the study was that the specification of a maximum time of 90 minutes for which a specimen should remain in the vehicle was critically reviewed, as the study demonstrated that there were considerable benefits from relaxing it.

Beltrami and Bodin (1974) addressed routing issues associated with waste collection. They comment that, in practice, the savings method is robust, despite the possibility to “lock onto locally optimal solutions” which can easily be seen to be inferior to solutions obtained by more naive methods. The practical issues addressed were the routing of hoist-compactor trucks (implemented in New York City) and the assignment of days to such routes, and the routing and dispatching of barges and tugboats in order to dispose of 10 000 tons of refuse each day at a landfill on Staten Island, a problem complicated by the tidal system and because a “customer” can be visited on more than one trip each day. The latter problem was tackled using the savings method in combination with a randomised search procedure in which the amount to be collected is randomly determined.

Atkinson (1990) reported an application carried out for the Inner London Education Authority (ILEA) to assist the scheduling of vehicles which deliver meals from kitchens to schools, using a heuristic based in part on the savings method and also incorporating a novel approach to the problem of ‘time window’ constraints on the collection and delivery
of meals. The problem addressed was large: there were 800 ILEA schools with their own kitchen, 190 of which supplied meals to schools without a kitchen. It is obvious that this is a time-critical application. The schedules produced for the first area to be studied showed a considerable improvement over the existing schedules, but this was not repeated when other areas were investigated. As a result, and also because of staff turnover, the life of the scheduling system proved to be rather short, and in any case ceased after the abolition of the ILEA in 1990, and the devolution of the school meals provision to thirteen local education authorities.

The logistic problems of ACSA, a company devoted to the production and distribution of frozen vegetables, based in Navarra (Spain), were addressed by Faulin (2003), using a so-called mixed algorithm (MIXALG). MIXALG uses both heuristic and exact subroutines to optimize the costs. As far as the routing aspect is concerned, a variant of the savings method (ALGACEA) is used, with impressive results, giving starting solutions that are improved by linear programming tools. The ALGACEA procedure chooses the customers whose saving has been selected by a Monte Carlo procedure. Once this heuristic part of the MIXALG method has been concluded, the routes found are improved by the application of a linear program to every Travelling Salesman Problem (TSP) in each route. These TSPs may be solved by exact methods because of the moderate size of the problem.

Pearn et al. (2004) investigated the wafer probing scheduling problem (WPSP), a variation of the parallel-machine scheduling problem, to sequence families of jobs on identical parallel machines with due date restrictions, by transforming it into a VRPTW (Pearn et al., 2002). Wafer probing is one of four main stages in the manufacture of integrated circuits. In their problem, the machine set-up time is sequentially dependent on the product types of the jobs processed on the machine, and the objective is to minimize the total machine workload without violating the machine capacity and job due date restrictions. The authors developed three modifications of existing savings algorithms, called the modified sequential, the compound matching based, and the modified compound matching-based savings algorithms, to solve the WPSP, which performed well according to the computational results.

Collect-and-place machines are one of the commonest types of placement machines in automated printed circuit board (PCB) assembly. Grunow et al. (2004) presented a three-stage heuristic solution approach to schedule the operations of this type of machinery, in which the second stage is modelled using adaptations of the savings method. In this stage, based on the assignment of component feeders to magazine positions determined in the first stage, the component placement sequence is determined. The component staging capacity of the revolver-type placement head creates a major constraint. As the total number of components to be placed on a board usually greatly exceeds the capacity of the revolver, several tours of the placement head have to be established, but the number of placement operations within a tour is restricted by the number of nozzles on the revolver. The placement-sequencing problem consists of three interrelated sub-problems, two of which, assigning components (placement operations) to the various tours and ordering the placement operations within each tour, can be considered as a VRP in micro-dimensions. The components required are collected from the warehouse (the magazine) and distributed to different customers (placement locations on the PCB). So VRP algorithms, specifically the
savings method, were adapted. In the third stage, local search principles were applied in order to improve the feeder assignment and the component placement sequence obtained. There is no indication given that these results were applied to assembling PCBs, rather numerical experiments were performed in order to compare the performance of the various savings-based heuristics under different experimental settings.

A decision-support system (DSS) was created for FRILAC Company (Navarra, Spain), to optimize its routes for delivering frozen food products on a network with 50 nodes (Faulin et al., 2005). Using the savings method, the DSS produces routes that minimize distances travelled, displays the final results on screen, creates reports for vehicle drivers, and estimates route costs. The choice of the savings method as the basis of the routing algorithm was not straightforward. After testing several algorithms the authors recommended the use of Fisher and Jaikumar’s algorithm (1981) to the logistics manager. He, however, favoured the use of the savings method on the basis of its simplicity, its intuitive nature, its low computation costs and the quality of its solutions. His arguments won the day. FRILAC successfully implemented the results, achieving an average cost saving of 10 percent.

An application of vehicle routing for Coca-Cola Enterprises (CCE), using ORTEC’s software that incorporates the savings method (Poot et al., 2002), was a recent Edelman competition finalist (Kant et al., 2008). CCE is the world’s largest marketer, producer, and distributor of Coca-Cola Company products (not only the well-known soft drinks, but also still and sparkling waters, juices, isotonics, teas, and energy, milk-based, and coffee-based drinks). In 2005, CCE distributed two billion cases (containing 42 billion bottles and cans), 20% of the Coca-Cola Company’s worldwide volume. Coca-Cola has outsourced its production and distribution to its bottling and distribution companies, of which CCE is the largest. CCE distributes syrup from the Coca-Cola plants to 64 bottling plants; it distributes bottled and canned beverages from the bottling plants to the distribution centres, and from the distribution centres to the final retail outlets. The OR study planned the distribution from over 430 distribution centres to 2.4 million outlets. The CCE fleet is the second largest in the United States after that of the US Postal Service: over 300 CCE dispatchers use this software daily to plan the routes of approximately 10,000 trucks. CCE realised an annual cost saving of $45 million and major improvements in customer service. This approach has been so successful that Coca-Cola has extended it beyond CCE to other Coca-Cola bottling companies and beer distributors.

7 Software for the savings method

It has long been widely assumed that the savings method is incorporated in most commercial software. As long ago as 1972 McDonald expressed surprise that the algorithms in 10 out of 12 commercially available vehicle scheduling packages evaluated in a 1969 survey (produced by the National Computer Centre in Manchester, UK) were based on the savings method (McDonald, 1972). In the same year, Wren and Holliday (1972) commented that “the method of Clarke and Wright and variations on it are commonly used in computer programs marketed for the problem.” Christofides et al. (1979) state that the savings method is “without doubt the most widely known heuristic and has formed
the basis of many programs including the IBM-VSPX package.” Baker (2002) reported that Truckstops®, the software available from Microanalytics, “uses a method based on the generalised assignment heuristic of Fisher and Jaikumar (1981) to seed routes and completes the route structure with a variant of the savings method.”

An email enquiry by the author to the sixteen vendors of vehicle routing software listed in the most recent Vehicle Routing Software Survey published by ORMS Today (Hall and Partyka, 2008) elicited eight responses. All but one specifically indicated that the savings method was incorporated into their algorithms to a greater or lesser extent.

By far the fullest account in the literature of the algorithmic component of commercial software is that concerning SHORTREC Distriplanner® (SDP), a commercial vehicle routing system, which ORTEC Consultants BV, a Dutch consultancy firm specialising in applied operations research, has sold since the mid-1980s to about 300 clients, mostly in Europe, Asia and the USA (Poot et al., 2002). Their experience was that real problems often “involve several non-standard constraints, some of which are never mentioned in the literature, while others are only analysed in isolation, i.e. not in combination with other non-standard constraints.” Adaptations of the savings method to incorporate such constraints have been implemented in SDP. Besides the traditional quality measures such as total distance travelled and total workload, the routing plans produced by the savings algorithm were compared with those produced by a sequential insertion algorithm, using real-life data, according to non-standard quality measures that help to evaluate the ‘visual attractiveness’ of the plan. Computational results show that, in general, the savings-based algorithm not only performs better with respect to these non-standard quality measures, but also with respect to the traditional measures. Currently, the savings method is not used by default, but only when dealing with few restrictions in a standard distribution environment. When there are many restrictions then an “insertion based” optimiser is used.

The main idea behind their savings-based algorithm is still to merge two trips with the highest saving, but adjustments are necessary to be able to cope with the various constraints. For a fixed and heterogeneous vehicle fleet, each time two trips are merged, the most appropriate vehicle has to be selected. At each iteration of the algorithm, the best possible combination of two trips is selected, but with a heterogeneous fleet it is frequently too time consuming to check the feasibility of each combination exactly, because this depends on the vehicle to which the merged route is assigned. Therefore, quick tests are carried out early in each iteration to eliminate combinations from further consideration. Only at a later stage are more elaborate feasibility tests carried out for relatively few combinations.

Customers in an area are allocated a region number. The second adjustment consists of a pre-processing step that has been added to reduce the size of the problem instance (the initial number of trips). So the customers are divided into groups and in each group as many customers as possible are joined together. Two customers are placed in the same group if, and only if, their postal codes match, their service intervals overlap, their region numbers match, and their feasible vehicle types match.

Since the customers have service intervals, unnecessary waiting time has been eliminated. Waiting time will normally not decrease in the next iteration of the algorithm (most
customers will keep the same predecessor and successor). Therefore, as a third adjustment, a maximum waiting time constraint is included. Only combinations of trips for which the total waiting time of the combined trip is less than a given value are permitted. Fourth, unnecessary waiting time in the combined trip is reduced as much as possible by applying a local search improvement method.

As a last adjustment, since some drivers are familiar with a particular region, it is preferable that these drivers’ vehicles serve the customers in that region. This is achieved by also assigning vehicles a region number. A region factor and a group bonus is included in the function that calculates the savings, ensuring that the algorithm is more likely to assign customers with a specific region number to a vehicle with the same region number and customers in the same group are more likely to be scheduled after one another.

8 Conclusions

This paper has sought to give insight into the origins of the savings method for the vehicle routing problem forty five years ago. Since then, several variations to the basic savings formula have been proposed and other improvements, mainly to improve the speed of computation, suggested. Furthermore, many extensions of the basic vehicle routing problem, which allows for constraints such as length of routes and capacity of vehicles, have been investigated using adaptations of the savings method, alongside many other approaches. In addition to the vast amount of research using the savings method, many practical problems have utilised the method, and it is a key component of the majority of the commercial routing packages. It is certainly true that the savings method has had a long and rich life, and it is apparent that this assessment is of an unfinished history, with the savings method continuing to make contributions to vehicle routing, and indeed other applications, as it approaches and passes its half century.

References


