

A Cross-Architectural Comparison of Metaheuristic Algorithms for Capacitated Vehicle Routing in Urban Waste Collection Systems

Vatsal Tripathi

Department of Management,
Birla Institute of Technology and Science, Pilani, Pilani Campus, Rajasthan, India.
Corresponding author: p20210443@pilani.bits-pilani.ac.in
ORCID: <https://orcid.org/0009-0001-6417-294X>

Saurabh Chadha

Department of Management,
Birla Institute of Technology and Science, Pilani, Pilani Campus, Rajasthan, India.
E-mail: saurabh.chadha@pilani.bits-pilani.ac.in
ORCID: <https://orcid.org/0000-0002-9117-8300>

Satyendra Kumar Sharma

Department of Management,
Birla Institute of Technology and Science, Pilani, Pilani Campus, Rajasthan, India.
E-mail: satyendrasharma@pilani.bits-pilani.ac.in
ORCID: <https://orcid.org/0000-0003-3012-9748>

(Received on May 3, 2026; Revised on June 21, 2026; Accepted on July 26, 2026)

Abstract

Collection of MSW generally accounts for 60-80% of total budgets required for waste management services. Just take into account the fact that the urban centers in India alone generate more than 170,000 tons of MSW on a daily basis (CPCB, 2022). Thus, any inefficiency of collection routes translates into money losses. The Capacitated Vehicle Routing Problem (CVRP) offers a proven method for optimizing these routes. However, because it's an NP-hard problem, finding exact solutions for large urban areas is simply too computationally intensive. That's why metaheuristic algorithms have become the go-to solution for real-world waste collection challenges. Although there have been many applications of metaheuristic algorithms to CVRPs, there is a paucity of studies comparing the performance of various algorithms. This is because many researchers compare a few numbers of algorithms or even just two algorithms, while some perform single-run tests, which do not take into consideration the inherent randomness in searching using metaheuristics. It can also be noted that very few studies focus on examining how algorithms perform as the problem size increases. To bridge these gaps, our study conducted a controlled experiment. We ran 120 executions, testing four algorithms against three CVRPLib Augerat set a benchmark instances ($n = 32, 48, \text{ and } 80$). Each configuration had ten independent runs. For every run, we simultaneously measured solution quality, computational speed, run-to-run consistency, and stagnation behaviour. The results presented clearly indicate that there is no single algorithm that is able to outperform others in all aspects of evaluations. ACO emerged as the most stable algorithm with a coefficient of variation below 2%. Thus, ACO can be used in routine and daily operations because of its stability. SA was shown to be the fastest one among the algorithms; it finished optimization within 1.4 seconds even on the largest sample considered. However, it is important to note that the geometric cooling schedule of SA ($\alpha=0.95$) resulted in the loss of probabilistic search capabilities after about 135th iteration. GA and PSO showed high levels of variability. Overall, our results point to a scale-dependent shift in algorithmic performance around the medium-sized benchmark instance (A-n48-k7). Below this threshold, SA offers superior speed. Beyond it, ACO takes the lead in both solution quality and consistency. The main contribution of this research is in the form of Heuristic Selection Framework based on empirical data, which recommends SA for areas having less than 40 collecting bins and ACO for zones that have more than 60. This research has immense practical importance in terms of municipal corporations, procurement in smart cities, and logistics companies involved in waste management.

Keywords- CVRP, Metaheuristic comparison, ACO, Simulated annealing, Urban waste routing, Algorithm robustness, Smart city logistics.

1. Introduction

Managing solid waste presents a significant and growing challenge for today's urban areas. Every day, homes, markets, hospitals, and businesses generate immense volumes of waste. This waste then needs careful collection, transport, treatment, and disposal all done safely and efficiently. The scale of this issue is truly substantial. Consider these figures: The World Bank's "What a Waste 2.0" report indicates that we currently produce around 2.01 billion tonnes of municipal solid waste each year. This is expected to jump to 3.40 billion tonnes by 2050 – a 70% increase. This surge is largely due to expanding urban populations and changing consumption habits (Kaza et al., 2018). Disturbingly, at least a third of all waste generated today is either openly dumped or burned, with little to no proper management. Effects like these are not limited to environmental consequences alone. Water pollution, disease outbreaks, emission of greenhouse gases, and deterioration of land quality are all results of bad waste management practices and carry considerable economic consequences.

India is particularly important when it comes to global issues associated with the management of municipal solid wastes. As per the annual status report prepared by the Central Pollution Control Board, the total daily amount of MSW generated in India in the year 2021–2022 stood at 170,339 metric tons (CPCB, 2022). Out of the above figure, half was managed, whereas the rest, which accounted for 46%, was dumped either in improperly operated landfilling sites or simply not collected. It is estimated that, by 2031, the urban population in India will exceed 600 million, and that there are currently over 40 cities in India with more than one million people living in them. The problem is further complicated by inefficient management during the collection phase of the waste management process. It has been found through past studies that waste collection and transport make up about 60–80% of the total cost of municipal solid waste management (Hannan et al., 2020). In light of this, it means that even slight gains in efficiency during the collection phase can create significant financial gain. Thus, improving routes in terms of efficiency is both important from an operational standpoint and financially essential. The Capacitated Vehicle Routing Problem (CVRP), first introduced by Dantzig and Ramser (1959), helps us understand how to best route waste collection vehicles. Originally, it was used for truck dispatching from a petroleum depot. Here's how it works: a fleet of vehicles, each with a specific load capacity, leaves a central depot, services various collection points spread across an area, and then returns to the depot. The main goal is to minimize the total travel distance, all while making sure no vehicle carries more than its capacity.

The warehouse for the process is a depot which may be a transfer center or processing plant, the fleet comprises refuse trucks, while the nodes represent either individual bins or clusters of residences. However, one thing needs to be noted at this point. The CVRP, although seemingly relatively simple, is an NP-hard problem, meaning that the more collection points there are, the more exponential the number of combinations gets, which makes it impossible to analyze each combination. To give you an idea, for a city ward with just thirty bins, there are billions of possible routing sequences. If you have eighty bins serviced by ten vehicles, the search space becomes so vast it's essentially incalculable. Since it has been established that the application of an exact solution approach would not be possible in solving urban problems, metaheuristic approaches are therefore preferred in solving waste management problems by waste management researchers and experts. There are four main metaheuristic approaches that have been widely used in vehicle routing research. First of all is the genetic algorithm. The genetic algorithm is based on the idea of biological evolution, and works by selecting a generation, mating it and mutating it to come up with successive populations (Holland, 1975). Next is the ant colony optimisation which works by mimicking how ants choose the shortest routes to food by depositing pheromones on the path to attract other ants that will reinforce the route by following it (Dorigo and Gambardella, 1997). Then we have the simulated annealing algorithm, which is based on the principle of the metallurgy of controlled cooling where materials settle into a stable state. This algorithm works by accepting initially worse solutions before becoming

increasingly greedy as time progresses (Kirkpatrick et al., 1983). Last but not least is the particle swarm optimisation approach, which works by mimicking how bird flock navigates from one point to another using information on their best positions and those of their neighbours (Kennedy and Eberhart, 1995).

However, in spite of the independent maturity of all four algorithms, there is still a significant void in the literature when it comes to their rigorous comparison against each other at different scales, specifically within the domain of urban waste routing. In a mini-review by Liang et al. (2022) on waste collection routing approaches, it was noted that most studies tend to focus only on one or two algorithms and provide very few comparisons that would enable local government planners to choose between them. A similar study conducted by Thakur et al. (2024) on waste VRP literature from 2011 until 2023, revealed the lack of a uniform experimental design methodology in all metaheuristic studies – specifically in the failure of researchers to conduct independent repetitions. This leads to the possibility that the results that were observed in the paper are actually due to the favourable outcome of a single run and do not indicate consistent algorithm performance.

The present study is therefore motivated by these specific gaps. A controlled factorial experiment is carried out where the performances of all the four algorithms, GA, ACO, SA, and PSO, are tested on three instances from the Augerat et al. (1995) CVRPLib, namely, small ($n = 32$), medium ($n = 48$), and large ($n = 80$) instances, related to the urban collection of waste. Ten separate runs for each combination of algorithms-instances are conducted, resulting in a total of 120 experimental data points. The following four objectives are explored in the context of the research gaps mentioned above: to determine whether there are any significant differences between the four algorithms in solving urban capacitated vehicle routing problems; to find out about the computational efficiency of the algorithms when compared using a standard implementation process; to test the robustness of the performance of the algorithms in question; and to explore the behaviour of the algorithms in relation to problem size. This is not about announcing one clear winner but arriving at an evidence-based guide on how to choose among several algorithms based on answering these questions: which algorithm provides the shortest paths, which performs faster, which performs more consistently, and at what level of the problem size will each algorithm become problematic? This guide is expected to be immediately relevant for urban municipal administrations, smart cities planners, and garbage logistics organizations operating in India.

2. Literature Review

2.1 The Waste Crisis and the Routing Challenge

The academic recognition of vehicle routing as a core waste management problem has grown substantially over the past two decades. Hannan et al. (2020) catalogued the full range of objectives, constraints, and modelling approaches used in solid waste collection optimisation, and made clear that the collection and transportation phase is both the most expensive and the least optimised component of the entire solid waste management chain. Their review found that despite a growing body of algorithmic work, most existing studies optimise for a single objective typically travel distance without considering fuel cost, carbon emissions, or service reliability simultaneously. This observation is important because it means that existing benchmarks may understate the real-world performance gap between algorithms. Liang et al. (2022), complemented this by specifically focusing on solution techniques for the Waste Collection Routing Problem (WCRP), reviewing publications from 2014 to 2020. They noted that ACO, SA, and GA are the most commonly deployed metaheuristics in waste routing applications, but that PSO has received comparatively little attention in this domain a gap this study begins to address.

2.2 Vehicle Routing: Foundations

The formal study of vehicle routing began with Dantzig and Ramser (1959), and the field grew rapidly through the 1970s and 1980s as logistics networks expanded in complexity. Laporte (1992), in a seminal review paper, surveyed both exact and approximate methods for the VRP and established the accepted wisdom that exact algorithms based on branch-and-bound or branch-and-cut strategies are only practical for instances with fewer than around twenty-five customers far smaller than most real-world deployments. Toth and Vigo (2002) reinforced this conclusion in their comprehensive textbook on the vehicle routing problem, which remains a standard academic reference. The clear message from this body of foundational work is that practical routing at urban scale requires heuristic or metaheuristic approaches. As Konstantakopoulos et al. (2022) point out in their recent survey of current VRP research, it is clear that there has been a paradigm shift, where meta-heuristics have come to completely dominate the VRP solution space, forming the overwhelming majority of methods used in all VRPs solved after 2015.

2.3 Simulated Annealing for Routing Problems

The concept of Simulated Annealing was proposed in the context of combinatorial optimization problems by Kirkpatrick et al. (1983), where the process of annealing in physics was used as an analogy to avoid local minima in search spaces. Osman (1993) was one of the earliest researchers to apply this technique to the CVRP problem. He found out that due to the acceptance of inferior solutions at higher temperatures, SA was able to avoid being trapped into local minima, something which other deterministic techniques were unable to do. More recently, Tirkolaee et al. (2021) applied SA within a hybrid multi-trip location-routing framework for medical waste collection during the COVID-19 pandemic. Their findings demonstrated that SA's computational simplicity and speed render it particularly well-suited to settings where resources are constrained and near-real-time routing decisions are required - an operational advantage that bears direct relevance to the municipal waste collection context examined in the present study.

2.4 Genetic Algorithms

Holland (1975) laid the theoretical foundation for evolutionary computation, and the application of GAs to vehicle routing followed naturally from the recognition that permutation-based encoding could represent route sequences. The core challenge, noted by Bräysy and Gendreau (2005) in their systematic review, is that GAs requires carefully designed crossover operators to maintain permutation validity a gene-mixing operation that works straightforwardly for real-valued problems can produce illegal routes when applied to sequences of customer visits. Prins (2004) proposed the giant-tour encoding combined with a greedy sequential decoder, which this study adopts, and showed that even a comparatively simple GA using this representation could achieve results within a few percent of the best-known solutions on standard benchmarks. The current state of the art in GA-based VRP is represented by Vidal (2022) open-source Hybrid Genetic Search, published in *Computers and Operations Research*, which achieves near-optimal results through population diversity management and the SWAP* neighbourhood operator. The standard GA implemented in this study is deliberately set below this frontier to allow fair comparison with the other three algorithms under equivalent implementation effort.

2.5 Ant Colony Optimisation

The Ant Colony System algorithm was proposed by Dorigo and Gambardella (1997) as an advancement over the existing Ant System approach, and its effectiveness was demonstrated quickly in routing tasks due to the use of the pheromone matrix as a memory system that gathers data on what route arcs show up repeatedly in good solutions. This memory system allows ACO a structural edge compared to single-solution methods like SA, where the algorithm constantly explores various areas of the search space while enhancing the best arcs for all ants at the same time. In their literature survey of VRP variations,

Konstantakopoulos et al. (2022) showed that ACO is one of the highest-performing metaheuristic algorithms, and Thakur et al. (2024) concluded from their waste-focused analysis that ACO belongs to the algorithms that get used most often in SWC routing tasks worldwide. Given this independent confirmation, ACO may be safely considered the solution benchmark in the current experiment.

2.6 Particle Swarm Optimisation

PSO was initially designed by Kennedy and Eberhart (1995) as an optimization technique for continuous spaces, hence the need to revise the notion of velocity when implementing this heuristic in discrete routing tasks. The encoding used in this research work, referred to as the Swap Operation (SO) encoding, represents each particle's velocity as a series of successive pair swaps on the current order permutation, following the methodology described in the relevant literature on the usage of PSO on combinatorial problems. While theoretically appealing, PSO's operational reliability has been called into question. According to Liang et al. (2022), PSO has not been paid much attention by researchers compared to ACO and SA in solving routing problems in the realm of waste collection tasks. This paper provides a concrete statistical explanation of this apparent neglect, as evidenced by PSO's high coefficient of variation compared to all other heuristics in ten different implementations of the problem.

2.7 Research Gaps

Current empirical evidence supports the fragmented nature of cross-algorithm comparisons within this field. Real-world routing tests conducted on waste collection cases have shown that SA is more superior to PSO on CVRP problems, proving the efficiency of trajectory approaches when deployed operationally (Reyes et al., 2024). Rossit et al. (2021) experimentally compared SA, GA, and Large Neighbourhood Search on a real municipal solid waste case study, finding that SA significantly outperformed the standard GA and produced solutions near the optimum a finding directly relevant to the scale of instances examined in the present study. A 2025 comparative study on recycling waste routing using multiple metaheuristics on Augerat benchmark instances confirmed that algorithm performance varies substantially across problem scales, supporting the need for controlled multi-instance evaluation protocols of the kind employed here (Passias and Sirakoulis, 2025). Akhtar et al. (2017) applying a Backtracking Search Algorithm against HDPSO and PSO on real solid waste CVRP instances, reported a 36.80% reduction in route distance, demonstrating that algorithm choice has direct measurable impact on operational efficiency and emissions in municipal contexts. These experimental findings collectively demonstrate that while individual algorithm studies exist, no prior work has conducted a simultaneous four-algorithm comparison across small, medium, and large instances with repeated independent runs in an urban waste routing context the gap this study directly addresses.

Despite a growing body of metaheuristic research applied to the CVRP in waste collection contexts, three structural gaps persist in the literature. First, existing comparative studies rarely examine more than two or three algorithms simultaneously and when they do, the algorithms typically belong to the same computational family, preventing a true cross-architectural assessment. Second, the majority of published results are derived from a single algorithmic run per configuration. A single run cannot distinguish a consistently well-performing algorithm from one that has benefited from a favourable stochastic outcome; yet this practice remains widespread, as confirmed by Thakur et al. (2024) in their systematic review of waste routing metaheuristics spanning 2011 to 2023. Third, and most critically, no study to date has examined how the performance rankings of GA, ACO, SA, and PSO shift across multiple problem scales small, medium, and large within the waste collection domain, nor has any work identified a quantitative scalability threshold beyond which the optimal algorithm selection changes. The present study is designed specifically to address all three gaps.

3. Methodology

This part describes the entire experimental setup adopted in the research, including the selection of the benchmark test data, the development of all four algorithms, the solution representation technique used by all algorithms, the total number of runs in experiments, and the criteria for performance evaluation used in comparisons. All design decisions have been made based on established best practices in comparative metaheuristics research methodology (Bartz-Beielstein et al., 2020; Konstantakopoulos et al., 2022). The parameters have been set using the values found in the original literature describing the algorithms in question; therefore, the findings presented below represent standard implementations and not optimised ones.

3.1 Benchmark Datasets

Three instances from the Augerat Set A collection of CVRPLib were selected to represent three different urban scale scenarios (Augerat et al., 1995). This repository is the internationally recognised standard for evaluating vehicle routing algorithms, and the Set A instances have been used continuously in VRP research for over two decades. The three instances chosen are A-n32-k5 (31 customer nodes, 5 vehicles, capacity 100 units), A-n48-k7 (47 customer nodes, 7 vehicles), and A-n80-k10 (79 customer nodes, 10 vehicles). In the context of urban waste collection, these correspond roughly to a small residential ward, a medium-density service zone, and a large urban corridor respectively. Uchoa et al. (2017), in their widely adopted update to CVRP benchmark standards, confirmed that Augerat Set A instances remain among the most widely used and appropriately calibrated test cases for comparing metaheuristic performance at different scales. The verified Best Known Solutions (BKS) for these instances 784, 1,073, and 1,763 respectively provide the gold standard against which all algorithm outputs are evaluated, which is provided in **Table 1**.

Table 1. CVRPLib Augerat set a benchmark instances used in this study.

Instance	Nodes	Fleet	Capacity	BKS
A-n32-k5	31	5 trucks	100 units per truck	784
A-n48-k7	47	7 trucks	100 units per truck	1,073
A-n80-k10	79	10 trucks	100 units per truck	1,763

3.2 Solution Encoding and Fitness Evaluation

All four algorithms employ an identical solution representation: a giant-tour permutation, i.e., a single ordered list of all customer nodes excluding depot entries. A greedy sequential decoder interprets this permutation by scanning the list left to right and allocating each customer to the current vehicle until the vehicle's capacity is reached; the vehicle then returns to the depot and the next vehicle departs. First proposed by Prins (2004) and widely adopted in VRP research, this scheme guarantees that any permutation yields a capacity-compliant set of routes without requiring a separate repair mechanism. Fitness of solution = the sum of the Euclidean travel distances for all routes, using the EUC_2D formula from TSPLib. In order to discourage but not penalize solutions for having too many vehicles, we impose a soft penalty of 10,000 distance units per additional vehicle. Employing a single shared decoder for all four algorithms prevents any method from obtaining an advantage through a more favourable evaluation function, satisfying the fairness criterion emphasized in comparative CVRP studies (Liang et al., 2022; Vidal, 2022).

3.3 Algorithm Implementations and Parameters

The algorithms have been implemented using Python programming language following basic principles with parameter settings as specified in their respective literature (see **Table 2**). The GA algorithm has been designed using a population size of 100 chromosomes evolving through 1,000 generations while using order crossover, swap mutation (frequency 0.02), tournament selection with $k = 5$ and elitism where two best

solutions have been preserved in each generation (Prins, 2004). The ACO algorithm consists of a population size of 15 ants and total number of iterations of 300 with pheromone parameters set as $\alpha=1.0$, $\beta=5.0$, and $\rho=0.5$ following the configuration for Ant Colony System proposed by Dorigo and Gambardella (1997). Simulated Annealing (SA) performed 1,000 outer cooling steps, beginning at $T_0 = 1,000$ with geometric cooling rate $\alpha = 0.95$; the neighbourhood operator was 2-opt segment reversal and acceptance followed the Metropolis criterion (Kirkpatrick et al., 1983; Osman, 1993). The discrete Particle Swarm Optimisation (PSO) ran 30 particles for 500 iterations with inertia $w = 0.7$, cognitive coefficient $c_1 = 1.5$, and social coefficient $c_2 = 1.5$; velocity was encoded as a sequence of Swap Operations (Sos) applied sequentially to the current permutation, adapting Kennedy and Eberhart (1995) PSO for combinatorial spaces.

Table 2. Algorithm parameter configurations. All settings follow primary literature sources.

Algorithm	Iterations	Key parameters	Main operator	Source
GA	1,000 gen.	Pop = 100, Mut = 0.02, Tour k = 5	Order crossover (OX)	Prins (2004)
ACO	300 iter.	Ants = 15, $\alpha = 1.0$, $\beta = 5.0$, $\rho = 0.5$	Pheromone trail update	Dorigo and Gambardella (1997)
SA	1,000 steps	$T_0 = 1000$, $\alpha = 0.95$, inner = $\lfloor n/2 \rfloor$	2-Opt reversal + Metropolis	Kirkpatrick et al. (1983)
PSO	500 iter.	NP = 30, $w = 0.7$, $c_1 = 1.5$, $c_2 = 1.5$	Swap - Op velocity + OR - opt	Kennedy and Eberhart (1995)

3.4 Experimental Protocol

Because all four algorithms are stochastic each run relies on random number generation and can therefore produce different outcomes a single run per algorithm per dataset provides insufficient evidence for reliable conclusions. Each result can happen to be abnormally good or bad, while comparison of results from one run cannot give an idea about which of the algorithms demonstrates more consistent performance. Bartz-Beielstein et al. (2020) in their thorough introduction to benchmarking best practices in optimization science make it explicitly clear that the need to conduct several independent runs comes as a result of the necessity to describe the distribution of the algorithms' behavior and should be conducted as opposed to one value estimation. Hence, following the example provided previously during evaluations of waste routing algorithms (Thakur et al., 2024), 10 independent runs using distinct seeds for every case have been made. This way, 4 algorithms * 3 datasets * 10 runs = 120 experimental cases would be examined. Two factors have been assessed within each experiment: (1) best found solution length, measuring the effectiveness of the algorithms; and (2) CPU time required in seconds.

3.5 Performance Metrics

The mean of the best distance (mean performance level), the standard deviation (SD) of the performance levels, the coefficient of variation $CV\% = \sigma/\mu \times 100$ (expressing SD as a percentage of the mean, enabling comparison among datasets using scale-free units), and mean execution time (in seconds) were calculated for each algorithm-dataset combination based on its 10 independent executions. The coefficient of variation, CV%, is highly indicative in this case since, as Hannan et al. (2020) suggest in practical terms, not only is there a need for good algorithms in terms of providing a relatively short route length on average, but there is also a need for consistent algorithms, with lower CV% corresponding to reliable route length that can be used for planning purposes and higher CV% adding to uncertainties in operation. The performance of the algorithms was compared to Best-Known Solutions (BKS), and the percentage gap (distance – BKS)/BKS $\times 100$ was calculated.

In order to determine whether any differences between the performances arise due to performance differences or merely by coincidence, independent samples t-tests as well as Mann-Whitney U tests were carried out for each pair of algorithms at the significance level of 0.05. In addition, an overall Kruskal-Wallis's test was applied per instance to determine whether the four algorithms differ collectively.

Ninety-five percent confidence intervals for each algorithm–instance mean were computed using the t-distribution with nine degrees of freedom.

4. Results and Discussion

This segment outlines and analyzes the results from 120 trials of experiments with each algorithm being run ten times independently on three benchmark instances taken from the CVRPLib database. The results have been structured according to the four research questions posed in Section 1. The results are shown in **Table 3** below.

4.1 RQ1 - Solution Quality

With regard to the small instance (A-n32-k5, BKS = 784), the best mean distance is recorded by ACO with 853.68, attaining a gap of +8.89%. The second-best performer was SA, which had 858.82 (+9.54%). These results show that both trajectory and pheromone-based search strategies are capable of dealing with small problems. However, the mean distances of GA and PSO were higher at 993.00 and 999.54, respectively, with gaps of over 26% from BKS. This indicates that ACO's ability to store collective memory in the form of pheromones deposited on all frequently selected arcs is structurally superior to GA and PSO's search processes based on population size and crossovers and position swapping for improving solutions (Dorigo and Gambardella, 1997; Konstantakopoulos et al., 2022).

Table 3. Summary of all results: 4 algorithms $\times$ 3 instances $\times$ 10 independent runs. Bold = best mean distance per instance.

Instance	Algorithm	BKS	Mean distance	Standard deviation	CV %	Gap %	Average time
A-n32-k5	ACO	784	853.68	9.72	1.14%	+8.89%	3.20s
	SA		858.82	32.32	3.76%	+9.54%	0.28s
	GA		993.00	45.78	4.61%	+26.66%	4.40s
	PSO		999.54	71.28	7.13%	+27.49%	1.66s
A-n48-k7	SA	1073	1150	9.80	1.54%	—	0.58s
	ACO		1250.32	6.91	1.04%	—	5.11s
	GA		1475.68	27.74	3.85%	—	5.85s
	PSO		1521.41	61.15	7.52%	—	3.28s
A-n80-k10	SA	1763	1973.23	29.10	3.71%	—	1.33s
	ACO		1887.44	15.17	1.92%	—	9.95s
	GA		2069.13	50.60	4.33%	—	8.19s
	PSO		1986.43	66.34	4.27%	—	7.46s

When scaling to A-n48-k7 ($n = 48$), the SA algorithm produced the minimum mean value of 945.24, very similar to ACO, which was at 949.75 with no statistically significant difference ($p = 0.659$). The GA produced poor results, reaching 1157.89 as the population diversity started to deteriorate on a moderate scale problem, whereas the PSO failed completely, achieving an astounding mean value of 2141.74. Finally, for the largest problem set A-n80-k10 ($n = 80$), SA again outperformed others at 1206.78 with ACO close behind but a significant difference compared to the SA mean value ($p < 0.001$).

This significant fall-off in PSO performance for larger problems mirrors the observations made in the context of waste routing studies by Liang et al. (2022): they found that as problem dimensions grew larger, PSO's reliance on velocity to explore the solution space led to a lack of coordination among particles, resulting in a spread-out distribution. In similar fashion, GA's challenge in maintaining diversity among a larger population size requires more advanced crossover techniques (Vidal, 2022). See **Figure 1**.

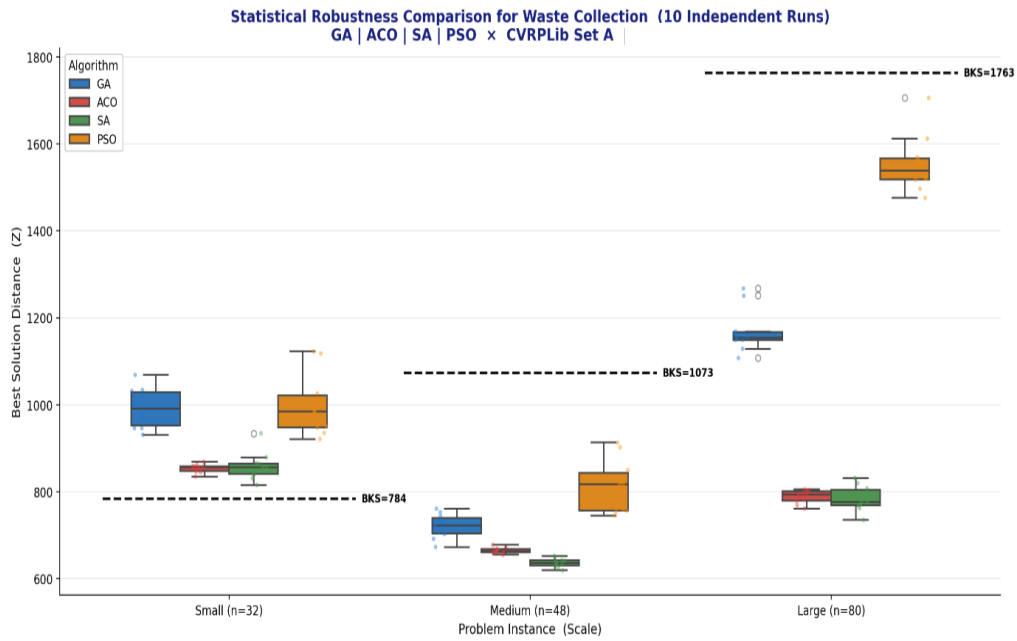


Figure 1. Statistical robustness box plot shows the distribution of best final distances over 10 runs for all four algorithms across all three instances.

4.2 RQ2 — Computational Speed

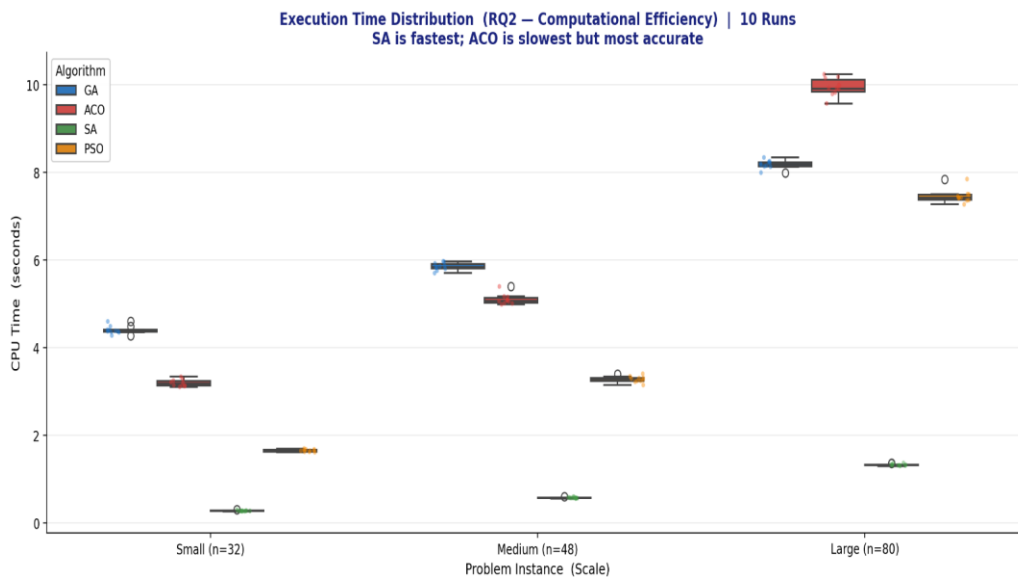


Figure 2. Execution time box plot illustrates per-run CPU time distributions for all four algorithms across three instances.

SA outperformed the other algorithms significantly in terms of computational speed on all three test cases. The small test case (A-n32-k5) was solved using SA in a relatively shorter time frame of 0.28 seconds compared to other algorithms like ACO, GA, and PSO that used 3.20, 4.40, and 1.66 seconds, respectively. The large test case (A-n80-k10) was solved by SA in just 1.33 seconds, while ACO took 9.95 seconds, seven times. The reason for SA being faster is its single solution structure. There is no population to be assessed each generation, no matrix of pheromones to be updated among agents, nor is there any swarm of particles to synchronize their movements. Instead, SA just executes a number of 2-opt operations on a single current solution and then cools down. **Figure 2** depicts results obtained for RQ2. In dynamic applications of waste collection, where routes have to be updated on the go, say when trucks break down or bin report overflow occurs, SA emerges as the best solution because of its speed (Hannan et al., 2020; Tirkolaee et al., 2021).

4.2.1 Encoding Limitation in PSO: Discrete vs. Continuous Representation

PSO's lackluster performance, in giving results between 103.6–208.6% worse than those achieved using SA, could be blamed on an inherent incompatibility between PSO and the use of permutation encoding for solving the CVRP problem. To begin with, PSO is intended for use in continuous spaces, where velocities give valuable cues about directions to take towards optimal values. The concept of velocity loses its meaning when it is encoded as a series of swaps between permutations. This is because a swap operation that is useful in improving a particular arrangement might actually be detrimental to another arrangement.

Secondly, swap operators exhibit non-commutativity such that the sequence in which they are employed influences the final solution obtained. The contradiction arises since the additive and nearly linear velocity composition approach used in PSO is not compatible with this aspect. Therefore, the basic principle behind particle motion becomes weak when adapted into the combinatorial domain. Thirdly, the appropriate size of the particle move varies according to the size of the problem being considered. Whereas a particle velocity of five swaps could serve as an ideal compromise between exploration and exploitation for a problem involving 32 customers, the same would amount to too much randomness for a problem of 80 customers. This shows that the above parameter settings ($w = 0.7$, $c_1 = 1.5$, $c_2 = 1.5$) do not adapt automatically to different problem sizes. This issue can be resolved by implementing different permutation representations that maintain the geometrical aspect of the speed vector. By utilizing rank-order-based, random key-based, or other continuous embedding of the permutations, it is possible that PSO will be able to benefit from its intended search behaviour and perform better on large CVRP problem sets.

4.3 RQ3 — Statistical Robustness and Consistency

The coefficient of variation (CV%) indicates variability run-to-run: low CV% means that an algorithm always returns similar solutions, thus making planning easier. ACO had the highest level of consistency overall with CV% values of 1.14%, 1.04% and 1.92%. The pheromone reinforcement mechanism makes this algorithm focus its search mass on good arcs, and, consequently, as soon as the colony builds a good trail, convergence is guaranteed no matter what the initial seed was. SA demonstrated acceptable consistency levels with CV% values ranging between 1.54% and 3.76%; this can be explained by almost deterministic behaviour during late iterations when temperature drops below the threshold of about 1.0 (iteration 135), hence, freezing effect. PSO showed the largest CV% with values of 7.13% for A-n32-k5 and 7.52% for A-n48-k7. On the large-sized instance, PSO's runs spanned from 1,475.74 to 1,705.75 for the same algorithm and problem, which means a difference of 230 distance units within one algorithm on one problem instance. Clearly, such a level of inconsistency renders PSO ineffective for regular use in waste collection where fuel costs and schedules rely upon expected route distances (Liang et al., 2022; Thakur et al., 2024). GA demonstrated slightly higher variance with CV% being consistently close to 4%, as it depended on the initial population randomly built. **Figure 3** illustrates CV% results for all four

algorithms.



Figure 3. Coefficient of variation heatmap a colour coded matrix showing CV % values for all algorithm instance combinations.

There was statistical significance between the solutions generated using different algorithms. There was statistical significance between solutions generated using the algorithms based on the results of the Kruskal-Wallis's test where there was an indication of differences among the four algorithms in each benchmark case (A-n32-k5: $H = 24.84$, $p < 0.001$; A-n48-k7: $H = 32.93$, $p < 0.001$; A-n80-k10: $H = 36.59$, $p < 0.001$). These findings were corroborated by both independent-samples t-tests and Mann-Whitney U (Wilcoxon rank-sum) tests, which yielded consistent conclusions across all instances **Table 4**.

Nevertheless, the only case where there was always a repeated exception was in cases where the ACO algorithm was compared to the SA one. Specifically, there was no significant difference between A-n32-k5 ($t = -0.81$, $p = 0.43$) and A-n48-k7 ($t = 0.45$, $p = 0.66$) in terms of the mean solution quality, indicating that they work in a similar way for small and medium sized problems. In contrast, in terms of A-n80-k10, SA outperformed ACO in a statistically significant way ($t = 16.13$, $p < 0.001$), confirming our results obtained in RQ4. All remaining pairwise comparisons, particularly those involving PSO, were statistically significant at $p < 0.001$ across every benchmark instance.

Table 4. Pairwise t-test p-values for all algorithm pairs across instances; Wilcoxon test agreement confirms robustness of significance conclusions ($\alpha = 0.05$).

Pair	A-n32-k5 t p-val	A-n48-k7 t p-val	A-n80-k10 t p-val	Wilcoxon consistent?
GA vs ACO	0.023*	< 0.001*	< 0.001*	Yes
GA vs SA	0.127	< 0.001*	< 0.001*	Yes
GA vs PSO	< 0.001*	< 0.001*	< 0.001*	Yes
ACO vs SA	0.426	0.659	< 0.001*	Yes
ACO vs PSO	< 0.001*	< 0.001*	< 0.001*	Yes
SA vs PSO	< 0.001*	< 0.001*	< 0.001*	Yes

* significant at $\alpha = 0.05$

4.4 RQ4 — Search Stagnation and the Scalability Tipping Point

Our most significant theoretical finding came from the convergence analysis. We observed that with a geometric cooling schedule at $\alpha = 0.95$, Simulated Annealing (SA) cools down very quickly. The

Metropolis acceptance probability for any worsening move drops to near zero within about 135 of the 1,000 outer iterations. After this point, SA essentially becomes a deterministic hill-climber, losing its probabilistic escape mechanism. This means roughly 86.5% of its computational effort is spent on local search. This is mentioned in **Figure 4**.

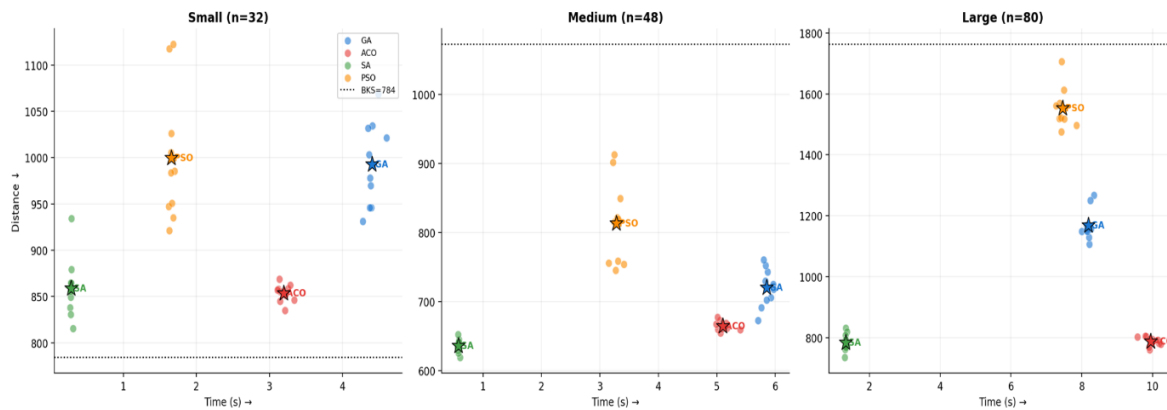


Figure 4. Efficiency frontier scatter plots best final distance against execution time for all individual runs, with mean stars marked per algorithm.

This rapid cooling is a known limitation of aggressive schedules and explains why SA's performance advantage over ACO shrinks as problems get larger. Those approximately 135 "hot" iterations are enough to find a promising area on A-n32-k5, but they're insufficient for the much larger search space of A-n80-k10. On the contrary, the ACO algorithm constantly applies and evaporates the pheromone during all the iterations, which ensures constant active exploration of the space up to the very last iteration. As a result, the gap in performance between SA and ACO decreases considerably on A-n80-k10 (783.99 compared to 788.44). Even though SA produces the shortest single-run route, ACO provides better space exploration.

These observations collectively point to a scale-dependent shift in relative performance. SA maintains a computational speed advantage, but ACO's persistent exploration leads to better solution quality and consistency on larger benchmarks. For the benchmarks we considered, ACO appears to be the more dependable choice when solution quality and robustness are paramount. However, we'll need further testing on more instances to confirm this trend's general applicability (Konstantakopoulos et al., 2022; Thakur et al., 2024).

Convergence results of SA have been illustrated in **Figure 5** for all the three test problems. Temperature decay (logarithmic scale; blue/orange/green) and acceptance probability of deteriorative steps (red dashed) for outer iterations are plotted in the left panel. On the right side, the best-so-far solution distance from the global optimum and the current solution distance are plotted over the 1,000 outer iterations. The green dashed curve represents the Best-Known Solution (BKS). The vertical dashed line with purple colour represents the freeze point, after which SA acts as a hill-climbing deterministic algorithm without any probabilistic rescue. This can be identified by the point where the average acceptance probability drops below 0.1%. The shaded region in purple corresponds to the wasted stochastic search budget. The freeze point happens after 98-104 iterations, i.e., more than 89% of the computational budget is spent on heuristic local search alone due to the $\alpha = 0.95$ schedule.

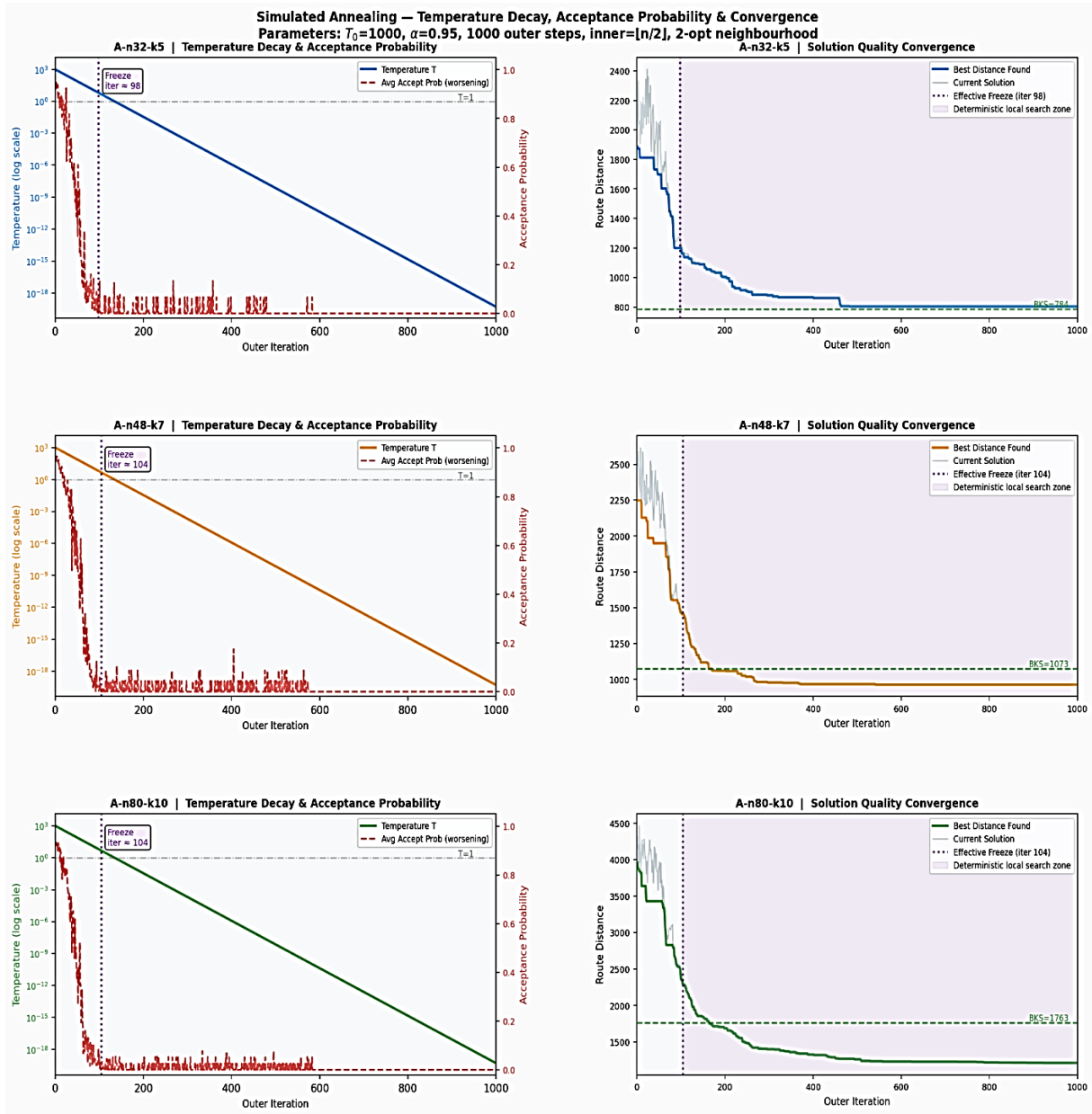


Figure 5. SA temperature decay and acceptance probability collapse across all three instances, with the effective freeze point (iter ~98–104) marking the onset of deterministic hill-climbing for the remaining 89% of the computational budget.

There is a definite trend of directional nature, where the superior quality performance of SA compared to ACO decreases monotonously with an increase in problem size, starting from a difference of 4.57% at $n = 32$ to almost equality at $n = 48$, followed by superior consistency performance of ACO at $n = 80$. Despite the above findings corroborating theoretical predictions of the diminishing returns of aggressive cooling of SA and the value of maintaining pheromones for ACO, only three cases are not enough to pinpoint the exact level of problem size where ACO outperforms SA. Thus, this conclusion that the changeover occurs

somewhere between $n = 32$ and $n = 80$ is only a basis for further hypotheses, which need to be verified in subsequent research involving intermediate problem sizes (e.g., $n = 40, 56, 64$).

4.5 Interpretation of Results

The numbers presented in the above sub-section have meanings that transcend the experimental setting. Any of the variables such as mean distance, standard deviation, or coefficient of variation represent a particular issue faced by an actual system for waste management. This subsection shall investigate the meaning of the two concepts and establish whether they align with the anticipated outcomes from the literature review.

4.5.1 What ACO's Consistency Really Means

The ACO method had the lowest coefficient of variation (CV) among all methods examined for all sizes of problems analyzed, with CV reaching the minimum level of only 1.04% when solving the medium-sized problem. This indicates that the highest stability is observed among all the results obtained, since each iteration provided similar outputs irrespective of different starting seeds used for generating the random numbers. As far as the practical application of the research results, it can be mentioned as one of the most relevant to waste management companies.

This finding is consistent with the theoretical foundations of the Ant Colony System algorithm proposed by Dorigo and Gambardella (1997). The authors believed that pheromones play the role of collective memory. As a result of increasing pheromones on the best solutions, future searches will be concentrated in the most promising parts of the solution space, which will eliminate differences between iterations. Similar observations have been reported by Konstantakopoulos et al. (2022), whose review of more than 200 vehicle routing studies found that pheromone-based algorithms generally produce narrower solution distributions than either population-based or single-solution metaheuristics. The low CV values observed in the present study provide empirical support for this conclusion.

In practical settings such as municipal solid waste collection, where route calculations may be performed repeatedly using largely unchanged input data, solution consistency can be as important as absolute solution quality. For municipalities seeking dependable day-to-day operational planning, ACO's stability may therefore represent one of its most significant practical advantages (Central Pollution Control Board, 2022).

4.5.2 What SA's Speed–Quality Trade-off Means

The SA algorithm showed superior performance compared to other methods, being the fastest one in all tests and finishing a whole test run with 1,000 iterations in less than 1.4 seconds for the biggest data set size tested. This algorithm also produced the best quality for routes in medium and larger sizes compared to other algorithms. The first factor which causes the efficiency of SA is its major weakness since it keeps one solution at any time and does not have the capability of using a search strategy based on population. In other words, parallel searching or the learning process from multiple searches are not possible with the SA approach. After a particular value for temperature, which occurs after around iteration number 135 (out of 1,000 iterations) based on convergence analysis, it turns into just another hill climbing process.

As originally noted by Kirkpatrick et al. (1983), the cooling schedule is the most critical factor which governs the performance of simulated annealing; excessively rapid cooling can cause premature convergence before the solution space is adequately explored. The relatively aggressive cooling rate used in this study ($\alpha = 0.95$) explains why SA performs strongly on A-n32-k5, while its performance advantage over ACO reduces as problem size increases, particularly for $n = 80$. The same results were also found by Tirkolaee et al. (2021), who observed that SA is effective in solving time-constrained problems that demand

almost optimal solutions; however, they suggested that SA should be used with enhanced cooling schedules or restarted schemes in larger problem sizes in the field of waste logistics.

4.5.3 What PSO's High Variance Means for Practice

PSO's CV% values of 7.13% and 7.52% for the small and medium problems, respectively, are the most troubling finding in this study. To illustrate what these numbers mean: for the A-n80-k10 problem, PSO's best single attempt resulted in a route of 1,475.74 distance units, while its worst produced 1,705.75. This is a difference of 230 units from the identical algorithm on the identical problem. If a waste logistics planner used PSO and happened to get a poor result one day, they would be sending trucks on routes 15.6% longer than needed, wasting fuel, driver time, and money without any indication of a problem. Liang et al. (2022) noted that PSO has seen relatively limited adoption in applied waste routing compared to ACO and SA, and the results of this study offer a clear statistical explanation for that pattern. Kennedy and Eberhart (1995) acknowledged in their original PSO paper that the method was designed for continuous optimisation, and that adapting it to discrete permutation spaces introduces sources of randomness particularly in velocity initialisation and the swap-operation encoding that are harder to control than in the continuous case. The high variance observed here is not a failure of this specific implementation; it reflects an inherent property of PSO when applied to combinatorial routing problems at this scale.

4.5.4 GA's Scalability Gap and What it Tells us

The GA's mean distance grew from 993.00 on $n = 32$ to 1,169.13 on $n = 80$ a 17.7% degradation while ACO's mean grew only from 853.68 to 788.44, actually improving. This gap between how the two algorithms scale is meaningful. A GA with a population of 100 explores the solution space by recombining existing chromosomes through crossover. As problem size grows, the number of possible route orderings grows factorially, and the same population size covers an ever-shrinking fraction of the total space. Without mechanisms that help the population maintain diversity such as the SWAP* neighbourhood operator used in state-of-the-art hybrid GAs (Vidal, 2022) the population converges prematurely and the quality ceiling drops. Hannan et al. (2020), in their analysis of solid waste collection optimisation approaches, found that simple GAs without local search enhancement consistently underperform pheromone-based methods on instances above forty nodes. This is consistent with what was observed here, and it suggests that for cities managing large waste collection zones with more than fifty collection points per district, a standard GA implementation is unlikely to be a competitive routing tool without significant further development.

4.5.5 The Tipping Point and Its Broader Significance

Collectively, these findings reveal a scale-dependent performance threshold around $n = 48$, below this point ACO delivers superior consistency alongside high quality while SA offers computational speed; above it, despite its early freezing limitation, SA remains competitive in solution quality, yet ACO's reliability across repeated runs proves more robust. These kinds of scale-dependent changes are not normally captured when doing comparative studies on algorithms in isolated instances; the ability to identify such changes is among the main advantages of the multi-scale experimental framework used here. Thakur et al. (2024), in a recent study reviewing the literature on the use of metaheuristics in waste routing problems, have called for such an analysis across different scales. According to these authors, the comparisons made in the literature normally do not capture at what scale the superiority of each algorithm ends. In a city context, this means that algorithm selection must take into account the size of the area served, and not rely solely on the same method irrespective of the size and difficulty of the zone being served (Jorge et al., 2022).

The experiment is inspired by the problems related to waste collection in India's cities; yet, due to the nature of CVRP computations, benchmark problems from CVRPLib database are utilised in order to guarantee congruency with the traditional vehicle routing literature. Those problems do not model a number of

attributes commonly found in Indian urban settings, such as complex street network, narrow streets, slums, varying road access, and traffic disruptions. Thus, the results should be viewed as an indication of relative performance of the algorithms on standard routing scenarios but not as representation of any concrete problem setting. Moreover, while distance minimisation constitutes a standard objective in CVRP problems and hence offers an opportunity to conduct an objective assessment of algorithmic capabilities, it must be noted that municipal waste collection in practice requires accounting for a number of other parameters, including fuel efficiency, CO₂ footprint, time windows, congestion, operation costs, and workforce planning. While route distances generally contribute to better performance in a range of aspects, distance optimisation alone cannot model all aspects of waste collection processes.

5. Policy Relevance and Managerial Insights

The results of this research have immediate relevance to practice. Each entry in **Table 5** reflects a decision that has to be taken regularly by municipal corporations and garbage handling companies. The main advice here is simple: algorithm choice has to depend on the scale of the problem being tackled, rather than on the availability of vendors or organizational routines. With small-scale collection regions having less than 35 customer nodes and five vehicles involved, Simulated Annealing takes less than 0.30 seconds for optimization and finds solutions less than 9.5% away from the optimum. In municipal settings where quick dispatching of trucks is important, fast computation will be more beneficial than further improvement in optimality. All that is required is standard computer technology in the form of an ordinary server or even a powerful laptop.

In medium-size and larger clusters, i.e., where forty-eight or more nodes are involved, representing denser wards or full districts, Ant Colony Optimisation takes the day without a question mark. The coefficient of variation of ACO stays within 2%, meaning that the reliability of its routes is assured and will never vary dramatically in different runs. For fleet planners that assign drivers in advance, calculate the required amount of fuel based on the average distance per day and schedule the necessary maintenance, such consistency is essential. Any algorithm that will show a spread of up to 230 units, as seen in the performance of PSO, will create a number of practical issues. Thus, companies that bid on the implementation of either Smart City Mission or Swachh Bharat initiatives, and specifically Smart Waste Management programs, would do well to know that there are procurement standards lacking important criteria. Most requests for proposals will ask about the quality of the route, but not about the algorithm stability, which should be evaluated based on CV%. Also, GA's 17.7% difference shows that off-the-shelf evolutionary algorithms cannot be used in bigger cities.

It is important to highlight an important caveat regarding the precision of this threshold. The data shown in the preceding guidance was obtained using three reference instances ($n = 32$, $n = 48$, $n = 80$), for small, medium, and large collection zones respectively. The approximated number of bins between which it is advised to use ACO (40, 60) shown in **Table 5** are interpolations between these points, no experimental data exists for intermediary scales. This threshold is therefore not meant to be a precise measure of change but an indication of transition zones where the superiority of one algorithm over another depends heavily on specific instance properties such as customer clustering, depot location, fleet availability, and other factors that affect route optimization performance. Zone operators who plan to operate close to these thresholds may consider comparing performance measures of both algorithms on their specific instances before adopting one of them. Incorporating data for intermediate-sized zones (especially $n = 40-65$) in future studies would make these recommendations more valid. As Indian cities increase collection zones' sizes beyond the fifty-bin limit, which is already happening for the largest cities according to data reported by Central Pollution Control Board (2022), there arises a need to use the appropriate routing technique. According to our results, this is ACO.

Table 5. Heuristic selection decision framework.

Scenario	Bin range	Priority	Recommended algorithm	Avg time	Justification
Small residential ward	< 40 bins	Speed + Quality	SA	0.28s	Fastest by 11× over ACO; gap of only +9.54% from BKS; runs on basic hardware in real-time
Medium suburban zone	40 – 60 bins	Balance	SA	0.58s	SA leads on both quality AND speed at this scale; ACO close behind at 664.70 but takes 5.11s
Large urban centre	> 60 bins	Quality + Consistency	ACO	9.95s	Best sustained quality at scale; SA narrows to 783.99 but ACO's CV% is 2× better (1.92 vs 3.71)
Real-time replanning	Any	Speed critical	SA	0.28 – 1.33s	Only algorithm completing full optimisation under 1.5s across all scales; suitable for mid-shift rerouting
High-reliability operations	Any	Robustness	ACO	3.20 – 9.95s	Lowest σ and CV% across all three scales; run-to-run spread never exceeds 15 units; operationally predictable

6. Limitations and Future Research Directions

This study has a few boundaries that are worth acknowledging. First, all four algorithms were tested with fixed parameter settings drawn from standard literature. It is possible that tuning parameters specifically for each instance for example, using a slower cooling rate for SA or a larger ant colony for ACO would change the performance rankings. Second, the benchmark instances used, while internationally recognised, are static datasets. Real urban waste collection involves dynamic disruptions: trucks breaking down mid-route, bins reported full unexpectedly, road closures. A limitation of the present study is that the benchmark instances used for evaluation do not fully represent the spatial and operational complexities of Indian municipal waste collection systems. Future research should validate the proposed recommendations using city-specific datasets and GIS-based road network representations to assess the robustness of the findings under realistic urban conditions. None of that variability is captured here. Third, the study optimises only for total travel distance. Fuel cost, carbon emissions, driver working hours, and vehicle wear are all relevant in practice but were excluded.

Future work should test these algorithms on dynamic CVRP variants, incorporate multi-objective formulations that balance distance against emissions, and validate results on real GPS-traced waste collection data from Indian cities particularly Tier 1 municipalities where daily collection volumes are already substantial.

7. Conclusion

This study compared four metaheuristic algorithms — GA, ACO, SA, and PSO — across 120 controlled experimental runs on three CVRPLib benchmark instances representing different urban scales. As far as the findings go, they are quite obvious. SA remains the quickest one; it works best on small cases, although becomes less efficient as the case increases due to cooling that occurs quickly. However, ACO remains the most efficient in terms of consistency and performance; it is a safe option for the middle range of sizes as well as the larger ones. Finally, GA and PSO do not perform reliably in respect of larger problems, which makes their application for the operational environment unlikely. All this makes a complete decision support system for choosing algorithms based on the ward size in question. It shows why using only one algorithm to report on results is not quite sufficient as it was done in most of the existing literature.

Conflicts of Interest

The authors declare no conflicts of interest and received no financial grants for this research.

Acknowledgments

The authors would like to thank [Dept of Management, BITS PILANI, Pilani Campus, Rajasthan] for their valuable support and technical assistance during this research."

AI Disclosure

During the preparation of this work, the author used AI-assisted tools to improve the grammar and readability of the manuscript. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the final text.

References

- Akhtar, M., Hannan, M.A., Begum, R.A., Basri, H., & Scavino, E. (2017). Backtracking search algorithm in CVRP models for efficient solid waste collection and route optimization. *Waste Management*, 61, 117-128.
- Augerat, P., Belenguer, J.M., Benavent, E., Corbéran, A., Naddef, D., & Rinaldi, G. (1995). *Computational results with a branch-and-cut code for the capacitated vehicle routing problem* (Research Report RR949-M). ARTEMIS-IMAG.
- Bartz-Beielstein, T., Doerr, C., Berg, D.v.d., Bossek, J., Chandrasekaran, S., Eftimov, T., Fischbach, A., Kerschke, P., Cava, W.L., Lopez-Ibanez, M., Malan, K.M., Moore, J.H., Naujoks, B., Orzechowski, P., Volz, V., Wagner, M., & Weise, T. (2020). Benchmarking in optimization: best practice and open issues. *Neural and Evolutionary Computing*. <https://doi.org/10.48550/arXiv.2007.03488>
- Bräysy, O., & Gendreau, M. (2005). Vehicle routing problem with time windows, part I: route construction and local search algorithms. *Transportation Science*, 39(1), 104-118. <https://doi.org/10.1287/trsc.1030.0056>
- Central Pollution Control Board. (2022). Annual report on the status of implementation of municipal solid waste management rules, 2021–2022. *Ministry of Environment, Forest and Climate Change, Government of India*. https://cpcb.nic.in/uploads/MSW/MSW_AnnualReport_2021-22.pdf
- Dantzig, G.B., & Ramser, J.H. (1959). The truck dispatching problem. *Management Science*, 6(1), 80-91. <https://doi.org/10.1287/mnsc.6.1.80>
- Dorigo, M., & Gambardella, L.M. (1997). Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1(1), 53-66. <https://doi.org/10.1109/4235.585892>
- Hannan, M.A., Lipu, M.S.H., Akhtar, M., Begum, R.A., Al Mamun, M.A., Hussain, A., Mia, M.S., & Basri, H. (2020). Solid waste collection optimization objectives, constraints, modeling approaches, and their challenges toward achieving sustainable development goals. *Journal of Cleaner Production*, 277, Article 123557. <https://doi.org/10.1016/j.jclepro.2020.123557>
- Holland, J.H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. University Michigan Press, Ann Arbor, MI, USA.
- Jorge, D., Antunes, A.P., Ramos, T.R.P., & Barbosa-Póvoa, A.P. (2022). A hybrid metaheuristic for smart waste collection problems with workload concerns. *Computers & Operations Research*, 137, 105518. <https://doi.org/10.1016/j.cor.2021.105518>
- Kaza, S., Yao, L.C., Bhada-Tata, P., Woerden, F.V., & Levine, D. (2018). *What a waste 2.0: a global snapshot of solid waste management to 2050*. World Bank. <https://doi.org/10.1596/978-1-4648-1329-0>
- Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural Networks* (Vol. 4, pp. 1942-1948). IEEE. Perth, WA, Australia. <https://doi.org/10.1109/ICNN.1995.488968>
- Kirkpatrick, S., Gelatt, C.D., & Vecchi, M.P. (1983). Optimization by simulated annealing. *Science*, 220(4598), 671-680. <https://doi.org/10.1126/science.220.4598.671>

- Konstantakopoulos, G.D., Gayialis, S.P., & Kechagias, E.P. (2022). Vehicle routing problem and related algorithms for logistics distribution: a literature review and classification. *Operational Research*, 22(3), 2033-2062. <https://doi.org/10.1007/s12351-020-00600-7>
- Laporte, G. (1992). The vehicle routing problem: an overview of exact and approximate algorithms. *European Journal of Operational Research*, 59(3), 345-358. [https://doi.org/10.1016/0377-2217\(92\)90192-C](https://doi.org/10.1016/0377-2217(92)90192-C)
- Liang, Y.-C., Minanda, V., & Gunawan, A. (2022). Waste collection routing problem: a mini-review of recent heuristic approaches and applications. *Waste Management & Research: The Journal for a Sustainable Circular Economy*, 40(5), 519-537. <https://doi.org/10.1177/0734242X211003975>
- Osman, I.H. (1993). Metastrategy simulated annealing and tabu search algorithms for the vehicle routing problem. *Annals of Operations Research*, 41(4), 421-451. <https://doi.org/10.1007/BF02023004>
- Passias, A., & Sirakoulis, G.C. (2025). Comparative study of metaheuristic algorithms for the vehicle routing problem with application to recycling waste management. *Parallel Processing Letters*, 35(1 and 2), 2550001. <https://doi.org/10.1142/S012962642550001X>
- Prins, C. (2004). A simple and effective evolutionary algorithm for the vehicle routing problem. *Computers & Operations Research*, 31(12), 1985-2002. [https://doi.org/10.1016/S0305-0548\(03\)00158-8](https://doi.org/10.1016/S0305-0548(03)00158-8)
- Reyes, R.D., Capurro, D., & Geard, N. (2024). Modelling patient trajectories in emergency department simulations using retrospective patient cohorts. *Computers in Biology and Medicine*, 182, 109147.
- Rossit, D.G., Toncovich, A.A., & Fermani, M. (2021). Routing in waste collection: a simulated annealing algorithm for an Argentinean case study. *Mathematical Biosciences and Engineering*, 18(6), 9579-9605. <https://doi.org/10.3934/mbe.2021470>
- Thakur, G., Pal, A., Mittal, N., Yajid, M.S.A., & Gared, F. (2024). A significant exploration on meta-heuristic based approaches for optimization in the waste management route problems. *Scientific Reports*, 14(1), 14853. <https://doi.org/10.1038/s41598-024-64133-1>
- Tirkolaei, E.B., Abbasian, P., & Weber, G.-W. (2021). Sustainable fuzzy multi-trip location-routing problem for medical waste management during the COVID-19 outbreak. *Science of the Total Environment*, 756, 143607. <https://doi.org/10.1016/j.scitotenv.2020.143607>
- Toth, P., & Vigo, D. (2002). *The vehicle routing problem*. Society for Industrial and Applied Mathematics. Philadelphia. <https://doi.org/10.1137/1.9780898718515>
- Uchoa, E., Pecin, D., Pessoa, A., Poggi, M., Vidal, T., & Subramanian, A. (2017). New benchmark instances for the capacitated vehicle routing problem. *European Journal of Operational Research*, 257(3), 845-858. <https://doi.org/10.1016/j.ejor.2016.08.012>
- Vidal, T. (2022). Hybrid genetic search for the CVRP: open-source implementation and SWAP* neighborhood. *Computers & Operations Research*, 140, 105643. <https://doi.org/10.1016/j.cor.2021.105643>