Green Vehicle Routing Optimization in Urban Traffic Congestion Environment Based On K-Means Clustering and Genetic Algorithm

Xiangyu Zha *

School of Mathematical Sciences and Statistics, Nanjing University of Information Science & Technology, Nanjing, China, 210044

* Corresponding Author Email: 15850335111@163.com

Abstract. As urban traffic congestion becomes increasingly severe, supply chain transportation faces the dual challenges of declining efficiency and mounting environmental burdens. To address this, this study constructs a Green Vehicle Routing Problem (GVRP) framework that integrates real-time traffic information. First, K-means clustering technology is employed to classify urban road networks by congestion levels, establishing a three-tier classification system of low, medium, and high congestion zones. Subsequently, polynomial regression methods are utilized to establish a quantitative relationship model between vehicle speed and carbon emission intensity. Based on this theoretical foundation, a multi-objective optimization framework is designed that comprehensively considers environmental costs and traffic impedance factors, with performance comparison tests conducted between genetic algorithms and classical shortest path algorithms. Experimental results demonstrate that genetic algorithms perform excellently when handling high-congestion road segments, significantly reducing carbon footprint while shortening transportation cycles. This research provides scientific basis and operational paradigms for enterprises to construct sustainable supply chain networks based on dynamic traffic information.

Keywords: Vehicle Routing Optimization, Genetic Algorithm, Shortest Path Algorithm, K-means Clustering Analysis.

1. Introduction

With the rapid development of modern cities and accelerated urbanization processes, traffic congestion has become a serious factor constraining the efficiency and reliability of supply chain transportation systems. Against the backdrop of rapid globalization and information technology development, the optimization design of supply chain networks holds significant importance for enhancing enterprise competitiveness and reducing operational costs [1]. However, urban traffic congestion not only leads to extended transportation times but, more importantly, frequent vehicle stops and re-accelerations increase fuel consumption, subsequently resulting in increased carbon emissions, thereby negatively impacting environmental sustainability [2]. With growing environmental awareness, incorporating carbon emission costs into supply chain total cost optimization frameworks has become an important topic in green supply chain research [3].

Addressing the aforementioned issues, this study proposes a traffic congestion level quantification method based on real-time traffic data and constructs a green supply chain route optimization model that integrates this factor. The research first employs K-means clustering algorithms to process actual traffic data, classifying road segments into three types: low congestion, medium congestion, and high congestion [4]. Subsequently, by collecting average vehicle speed data on roads under different congestion levels, polynomial regression analysis is applied to establish an estimation model between speed variations and carbon emissions [5]. This model reflects that carbon emissions are not solely determined by travel distance but are influenced by road congestion levels and driving speeds, exhibiting nonlinear variation characteristics.

Based on the above data, this paper establishes a Green Vehicle Routing Problem (GVRP) model that simultaneously considers carbon emission costs and congestion impacts, employing genetic algorithms (GA) and traditional shortest path algorithms respectively for route searching and comparative analysis [6]. Empirical results demonstrate that compared to shortest path algorithms that only consider geographical distance, methods based on K-means clustering and GA optimization can significantly reduce total operating costs and carbon emissions [7]. Particularly in high-congestion areas, the GA method demonstrates distinct advantages, capable of simultaneously reducing carbon emissions and transportation time, thereby achieving unity between environmental friendliness and efficiency in supply chain routing [8].

This study not only provides theoretical contributions for utilizing real-time traffic data in environmentally friendly supply chain design but also offers practical guidance for urban logistics enterprises in formulating green delivery strategies under complex traffic environments [9]. Through integrating big data analysis and optimization technologies, this research demonstrates the potential for applying this method in actual operations, providing theoretical basis and decision support for enterprises to achieve green logistics and sustainable development in complex urban environments [10].

2. Model Construction and Algorithm Design

2.1. Traffic Congestion Pattern Recognition Based on K-means Clustering

To effectively integrate the dynamic and complex urban traffic conditions into the route planning model, this research first employs the K-means clustering algorithm to analyze the collected traffic data. As a widely applied unsupervised learning technique, the core principle of the K-means algorithm lies in partitioning objects in a dataset into K distinct clusters, such that the similarity among objects within the same cluster is maximized while the similarity between objects in different clusters is minimized. This algorithm can be represented as minimizing the following objective function:

$$J = \sum_{k=1}^{K} \sum_{x_i \in C_k} \| x_i - \mu_k \|^2 \tag{1}$$

In the application scenario of this research, the K-means algorithm is used to process traffic data containing 297 Beijing Road geographical location points obtained from Amap, including congestion indices and average vehicle speeds. The specific execution process of the algorithm follows standard steps: first, K data points are randomly selected as initial clustering centers; then, the distance from each data point to each clustering center is calculated, and the point is assigned to the cluster of the nearest center; subsequently, the mean of all data points within each cluster is recalculated, and this mean becomes the new clustering center; this assignment and update process is repeated iteratively until the clustering centers no longer undergo significant changes or the preset maximum number of iterations is reached. This research employs the silhouette coefficient method for decision-making.

The results of K-means clustering provide a crucial foundation for subsequent route optimization. By assigning each geographical location point to a specific congestion category, the model is able to link abstract traffic conditions with specific spatial locations. This enables the determination of a more realistic average vehicle travel speed when calculating route segments between any two points i and j, based on the congestion zone categories they belong to or traverse. This speed parameter subsequently becomes a key input for calculating carbon emissions and estimated travel time for that road segment, thereby quantifying the actual impact of traffic congestion and integrating it into the cost function of the GVRP model.

2.2. GVRP Route Optimization Based on Genetic Algorithm (GA)

The GVRP model proposed in this research involves finding a path with the lowest total cost in a complex network containing a fixed starting point, a fixed endpoint, and five candidate intermediate stations, which is a typical combinatorial optimization problem. Given the enormous solution space and the nonlinear characteristics of the objective function, traditional exact algorithms are difficult to solve within reasonable time. Therefore, this research chooses to adopt the Genetic Algorithm (GA), a heuristic optimization technique, to find high-quality near-optimal solutions.

GA maintains a population composed of potential solutions and simulates mechanisms such as natural selection, inheritance, and mutation, enabling the population to gradually evolve toward better solution regions during the iterative process. The core advantage of GA lies in its global search capability and robustness in handling complex, nonlinear problems, effectively avoiding local optimum traps.

In this research, the implementation of GA is specifically designed for the particular structure of the GVRP model:

Encoding: Each individual represents a complete W → Sa → Sb → Sc → Sd → Se → C path. Since the 5 intermediate stations need to be selected from the candidate set S without repetition, a chromosome can be a sequence containing 5 different station indices, where these indices correspond to specific stations in the candidate station set S, and their order represents the visiting sequence.

Initialization: At the beginning of the algorithm, an initial population needs to be created. This is accomplished by randomly generating a set of permutation-encoded chromosomes that satisfy the constraints. The size of the initial population needs to balance solution diversity and computational efficiency.

Fitness Evaluation: Each path scheme represented by a chromosome needs to be evaluated for its quality. The fitness function directly corresponds to the optimization objective of this research, namely minimizing total cost. Since GA typically maximizes fitness, the fitness value can be set as the reciprocal of total cost or a sufficiently large constant minus the total cost, so that paths with lower costs have higher fitness. The fitness value of each chromosome is defined as:

$$F(X) = C_{\max} - Z(X) \tag{2}$$

Selection: The selection operation simulates the principle of "survival of the fittest," tending to give individuals with higher fitness more opportunities to reproduce the next generation. This research adopts the roulette wheel selection method. The probability of an individual being selected is:

$$P(X_i) = \frac{F(X_i)}{\sum_{j=1}^{m} F(X_j)} \tag{3}$$

Crossover: The crossover operation generates new offspring individuals by exchanging partial genes of two parent individuals, serving as the main method for GA to explore new solution spaces. This research adopts the single-point crossover method, randomly selecting two parent chromosomes, randomly determining a crossover point position, then exchanging the gene segments after the crossover point between the two parents to generate two offspring chromosomes. The crossover operation is performed with a certain crossover rate.

Mutation: The mutation operation randomly changes certain gene values on individual chromosomes with a small probability, aiming to introduce new genetic material, increase population diversity, and prevent the algorithm from prematurely converging to local optimal solutions. This research adopts the random mutation method. In a selected chromosome, one or more gene positions are randomly chosen, and their values are randomly changed to another valid station index from the candidate set S. The mutation operation is performed with a certain mutation rate.

Iteration and Termination: Algorithm terminates after iterating for generations, or when the optimal solution shows no significant improvement for consecutive generations.

$$|F_{\text{best}}^{(g)} - F_{\text{best}}^{(g-T)}| < \epsilon \tag{4}$$

2.3. Shortest Path (SP) Algorithm as Baseline Comparison

To objectively evaluate the effectiveness of the GVRP model and GA optimization method proposed in this research, we introduce the shortest path algorithm as a baseline for comparison. The objective of the SP algorithm is to find the path with the shortest geographical distance between a given starting point and endpoint, which can be expressed as:

$$\min Z_{SP} = \sum_{(i,j) \in A} d_{ij} \cdot y_{ij} \tag{5}$$

3. Experimental Analysis and Results Discussion

3.1. Data Clustering Analysis

The traffic data for the GVRP model proposed in this study was collected through the Amap API on November 1, 2024, encompassing a total of 297 Beijing Road traffic congestion data points. Based on this data, congestion indices and average vehicle travel speeds for 297 road segments were obtained.

Through the K-means clustering algorithm, the 297 traffic data points in the study were divided into three congestion levels: low congestion (Ka), medium congestion (Kb), and high congestion (Kc). Figure 1 shows the geographical distribution after clustering, where low congestion areas are represented by red circular points (Ka), medium congestion areas by green square points (Kb), and high congestion areas by blue diamond points (Kc). From the distribution maps, it can be observed that different congestion level areas exhibit certain spatial clustering characteristics, which align with actual traffic conditions. Low congestion areas are mainly distributed in urban periphery and some main arterials, medium congestion areas are located in the urban middle ring and secondary roads, while high congestion areas are concentrated near the city center and transportation hubs.

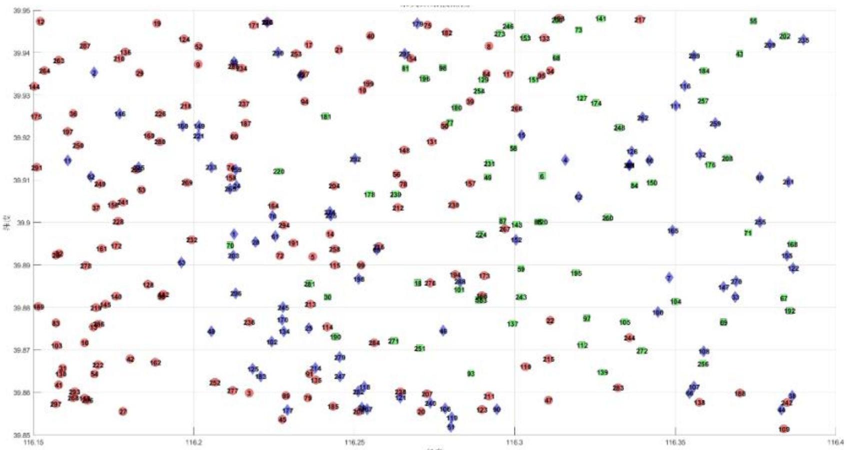

Figure 1 Clustering Results Scatter Plot

3.2. Carbon Emission Model Fitting Analysis

As figure 2, the GVRP model proposed in this research aims to dynamically plan supply chain routes and find paths that minimize total costs. This model considers traffic congestion factors; therefore, carbon emission costs generated by traffic congestion are also included in the total costs. In this study, carbon emission costs are parameterized through polynomial fitting to calculate carbon emissions at different speeds caused by traffic congestion. This curve exhibits typical U-shaped characteristics, indicating that vehicles produce lower carbon emissions at moderate speeds. When speeds fall below or exceed this optimal range, carbon emissions increase significantly.

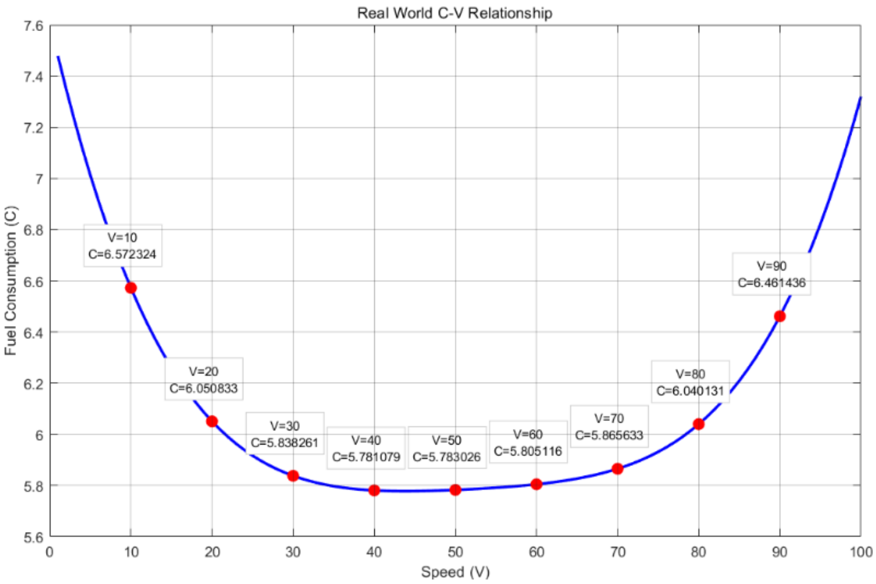

Figure 2 Carbon Emission Model Fitting plot

3.3. Route Optimization Results and Analysis

The GVRP model in this research aims to perform dynamic route planning and utilizes the Genetic Algorithm (GA) to minimize total cost. To observe the changes in total cost and carbon emission cost, we also calculated the total cost and carbon emission cost using the shortest path as a baseline. To increase the reliability of the results, in each scenario, we selected 3 different starting points, labeled as Ka_1, Ka_2, Ka_3, Ka_4, Ka_5 (Scenario 1), Kb_1, Kb_2, Kb_3, Kb_4, Kb_5 (Scenario 2), and Kc_1, Kc_2, Kc_3, Kc_4, Kc_5 (Scenario 3), respectively. Thus, we obtained a total of 45 different route optimization results.

3.3.1 GA Optimization Results Analysis

Table 1 GA Optimization Route Results

GA	Starting Point	Ending point	Route	Distance	Cost	Carbon Emission Cost
Scenairo1	ka_1	ka	212-114-284-118-276-131-173	16.84306	8345.036	134.7445
		kb	212-191-1-14-214-284-101	14.94643	8575.729	124.8729
		kc	212-14-214-25-114-102-48	10.15837	8789.843	86.64523
	ka_2	ka	41-206-96-14-245-284-173	17.28379	8616.637	118.5715
		kb	41-96-1-14-131-276-101	16.76223	8690.567	128.5025
		kc	41-27-206-13-140-96-48	8.583869	9166.572	54.72554
	ka_3	ka	182-131-148-14-214-276-173	15.35417	8957.928	122.8333
		kb	182-98-131-14-284-276-101	12.53537	8769.383	107.6237
		kc	182-98-131-114-284-102-48	14.1429	9089.38	122.1729

	ka_4	ka	10-253-227-292-14-131-173	14.6584	8612.895	122.4892
		kb	10-131-148-14-284-276-101	12.02252	8714.666	97.75035
		kc	10-292-14-1-96-206-48	16.07573	9017.293	125.4951
	ka_5	ka	162-206-27-96-14-131-173	18.40018	8338.083	125.198
		kb	162-96-206-114-214-284-101	17.21444	8251.002	119.5096
		kc	162-96-245-214-114-102-48	11.05541	8741.64	91.84741
Scenairo2	kb_1	ka	174-131-148-292-14-98-157	17.04483	8556.377	164.7569
		kb	174-127-98-246-8-129-254	8.167493	8636.585	95.55862
		kc	174-127-98-131-14-96-48	16.13548	8622.192	140.9617
	kb_2	ka	176-126-152-284-14-131-157	18.78403	8412.336	154.4774
		kb	176-126-98-129-131-152-254	18.23378	8507.009	177.7394
		kc	176-126-20-284-114-102-48	17.57433	8570.319	148.0589
	kb_3	ka	190-114-96-233-1-14-157	16.79052	8821.558	129.7425
		kb	190-114-284-20-14-131-254	14.52452	8772.684	120.5148
		kc	190-114-245-1-96-206-48	13.57103	8935.09	107.1447
	kb_4	ka	200-73-8-98-129-131-157	9.123284	8649.942	92.99329
		kb	200-73-8-98-131-129-254	8.801022	8729.826	91.17601
		kc	200-98-131-1-70-96-48	15.19752	8829.208	144.3541
	kb_5	ka	93-123-20-284-114-131-157	13.19972	8683.829	107.5806
		kb	93-284-114-214-14-131-254	14.1066	8782.062	121.8982
		kc	93-284-114-118-214-102-48	10.23203	8908.585	95.9891
Scenairo3	kc_1	ka	160-131-14-25-114-118-284	16.29408	8331.311	143.6111
		kb	160-233-96-114-284-33-192	21.91028	9051.385	174.925
		kc	160-233-36-136-9-284-167	19.74216	8864.379	135.0922

	kc_2	ka	122-155-242-33-147-152-284	18.00026	7829.542	170.7265
		kb	122-155-33-147-38-242-192	8.540164	8503.442	85.40164
		kc	122-33-123-20-284-118-167	14.64464	8362.253	121.7624
	kc_3	ka	259-126-131-292-14-114-284	15.56933	7853.414	133.3881
		kb	259-235-126-33-147-155-192	17.04751	8777.255	162.527
		kc	259-126-98-131-276-284-167	17.37827	8427.626	158.5719
	kc_4	ka	119-20-118-214-102-114-284	8.053503	7971.217	71.57909
		kb	119-20-284-33-147-242-192	18.98615	8812.825	166.1261
		kc	119-20-284-118-114-214-167	8.908562	8485.794	76.42469
	kc_5	ka	46-276-14-214-114-118-284	11.92215	8357.645	95.37718
		kb	46-284-118-20-242-33-192	18.98756	9062.019	164.0226
		kc	46-284-214-114-118-205-167	7.428066	8765.275	60.76548

In Scenario 1, departing from low congestion area Ka, which includes fifteen routes, the total costs are distributed between approximately 8,251 yuan to 9,167 yuan, while carbon emission costs range from 54.73 yuan to 134.74 yuan. The data shows that even when the starting point has favorable traffic conditions, the routes selected by the GA algorithm do not simply pursue the shortest distance. For example, the route from starting point ka_2 to endpoint kc has the shortest travel distance and the lowest carbon emission cost, but the highest total cost in this scenario, reflecting the significant impact of fixed costs or processing costs associated with other nodes on the route on the total cost. In contrast, the route from starting point ka_5 to endpoint kb, although longer in distance, achieves the lowest total cost in this scenario, demonstrating the algorithm's ability to balance multiple cost factors.

In Scenario 2, departing from medium congestion area Kb, the total costs of the fifteen routes range from approximately 8,412 yuan to 8,935 yuan, with carbon emission costs ranging from 91.18 yuan to 177.74 yuan. The overall costs in this scenario are generally slightly lower than those in Scenario 1. Particularly noteworthy is the route from starting point kb_4 to endpoint kb, which has the shortest total distance, near-lowest carbon emission cost, and moderate total cost. Meanwhile, the route from starting point kb_2 to endpoint kb, despite not having the longest total distance, generates the highest carbon emission cost in this scenario, highlighting that even when the starting point and endpoint have the same congestion level, differences in intermediate route selection have enormous impacts on carbon emissions, validating the importance of congestion conditions in internal route segments.

In Scenario 3, departing from high congestion area Kc, the total costs of the fifteen routes show a lower trend, ranging from approximately 7,830 yuan to 9,062 yuan, with carbon emission costs between 60.77 yuan and 174.93 yuan. This indicates that when transportation tasks begin in areas with severe traffic congestion, the GA algorithm has greater optimization potential. By intelligently selecting intermediate delivery stations, the algorithm can effectively avoid sustained high-congestion road segments, thereby significantly reducing overall operational costs. For example, routes from starting point kc_4 to endpoint ka and from starting point kc_5 to endpoint kc not only have relatively short total distances and low carbon emission costs, but also achieve lower total costs. In contrast, the route from starting point kc_1 to endpoint kb has the longest distance, with correspondingly high carbon emission costs and total costs.

3.3.2 Shortest Path Results Analysis

To evaluate the effectiveness of GA-optimized routes, we also calculated the shortest paths (SP) based on geographical distance as a comparison. Table 2 presents detailed information for the 45 shortest paths.

Table 2 SP Optimization Route Results

GA	Starting Point	Ending point	Route	Distance	Cost	Carbon Emission Cost
Scenairo 1	ka_1	ka	212-276-101-57-163-166-173	4.233848	10897.89	87.12443
		kb	212-239-18-276-194-288-101	4.214117	10173.5	101.12
		kc	212-99-30-213-245-236-48	6.258424	11867.43	117.6764
	ka_2	ka	41-54-48-236-276-288-173	11.70399	10471.68	204.5464
		kb	41-54-42-162-102-25-101	11.00092	10101.56	222.6192
		kc	41-268-113-206-222-162-48	5.091267	10633.47	70.69046
	ka_3	ka	182-180-77-50-230-224-173	6.799295	11509.78	144.2471
		kb	182-180-77-230-194-288-101	6.828434	11127.38	146.9219
		kc	182-81-292-294-61-296-48	10.11742	12056.13	232.8971
	ka_4	ka	10-199-148-56-78-230-173	6.654255	10757.73	117.1149
		kb	10-148-230-224-194-288-101	6.614011	10598.78	129.4826
		kc	10-181-164-76-61-296-48	7.733425	12080.89	168.9374
	ka_5	ka	162-236-245-18-276-288-173	9.115295	10006.17	184.8487
		kb	162-48-236-245-18-276-101	8.465613	10314.02	185.9035
		kc	162-42-13-286-145-128-48	5.998561	11522.76	97.52849
Scenairo 2	kb_1	ka	174-95-151-266-231-49-157	5.203384	11287.82	116.0757
		kb	174-127-95-151-64-129-254	3.609039	11293.57	90.62259
		kc	174-230-216-65-245-236-48	11.92171	11615.43	243.0008
	kb_2	ka	176-150-88-201-6-49-157	6.645809	11455.85	186.654
		kb	176-262-174-127-151-117-254	6.822351	11200.11	165.6712
		kc	176-150-84-143-198-245-48	13.9575	11283.7	373.8706

	kb_3	ka	190-198-99-65-216-212-157	5.836147	12386.23	117.0885
		kb	190-65-216-212-50-39-254	7.527074	12587.76	156.2647
		kc	190-114-134-170-245-236-48	4.139874	11304.11	85.00636
	kb_4	ka	200-68-34-95-151-231-157	5.084664	11643.77	130.2606
		kb	200-133-153-273-64-129-254	3.372279	11083.47	87.27807
		kc	200-254-39-50-131-56-48	12.28837	11251.68	242.3237
	kb_5	ka	93-57-163-166-173-230-157	5.39729	11758.36	113.4963
		kb	93-57-166-173-224-157-254	7.482281	11219.04	170.8057
		kc	93-271-284-114-25-236-48	7.180377	10289.75	145.9531
Scenairo 3	kc_1	ka	160-158-24-61-281-30-284	7.656504	11732.24	161.6919
		kb	160-149-60-230-87-147-192	16.86241	10985.55	375.0208
		kc	160-233-72-213-114-279-167	8.920276	10762.79	169.6112
	kc_2	ka	122-104-100-22-137-271-284	11.38382	11318.64	266.2647
		kb	122-155-255-270-33-67-192	5.244233	10542.68	124.2559
		kc	122-270-104-100-105-93-167	11.93861	12061.39	288.0321
	kc_3	ka	259-201-62-85-152-288-284	10.76341	9520.845	233.3539
		kb	259-176-255-168-155-122-192	5.951939	10259.73	155.5093
		kc	259-86-59-163-46-251-167	11.96492	11990.87	300.2379
	kc_4	ka	119-106-240-121-238-271-284	3.538742	10867.01	75.55783
		kb	119-90-47-283-108-69-192	9.667755	12276.67	213.3528
		kc	119-106-240-20-238-26-167	3.050476	11165.07	58.93004
	kc_5	ka	46-207-240-20-238-271-284	4.915783	11047.29	95.04176
		kb	46-137-22-97-105-100-192	9.268021	12337.14	235.964
		kc	46-207-240-20-121-205-167	4.367526	10583.81	87.54166

By comparing the data in Table 1 with Table 2, it can be found that although the SP method almost always outperforms or equals the GA method in its sole optimization objective - total travel distance - the total costs and carbon emission costs it calculates are generally significantly higher than those of GA-optimized results. The total costs of the SP method are roughly distributed in a higher range of approximately 9,500 yuan to 12,600 yuan, while its carbon emission costs fluctuate dramatically, from a minimum of about 58.93 yuan (starting point kc_4 to endpoint kc) to a maximum of 375.02 yuan (starting point kc_1 to endpoint kb). Both cost indicators are significantly higher than the average and optimal levels of GA-optimized routes.

Particularly noteworthy is that the SP method generated multiple routes with extremely high carbon emission costs. For example, in Scenario 2, the SP route from starting point kb_2 to endpoint kc has a carbon emission cost as high as 373.87 yuan, which is almost two and a half times the 148.06 yuan carbon emission cost of the route found by GA for the same origin-destination pair. Similarly, in Scenario 3, the SP route from starting point kc_1 to endpoint kb also reaches an alarming carbon emission cost of 375.02 yuan, far exceeding the 174.93 yuan of the corresponding GA route. This phenomenon greatly reveals that the strategy of purely pursuing the shortest geographical distance, while ignoring actual traffic operating conditions, may lead to serious environmental negative externalities and economic losses.

From a route selection perspective, the SP method tends to select geographically most direct points for connection, which often means that routes will inevitably traverse urban central areas or known traffic congestion hotspots. Taking the route from starting point kb_2 to endpoint kc in Scenario 2 as an example, the SP-selected route "176-150-84-143-198-245-48" has a total distance of only 13.96 kilometers, which is relatively short among all SP routes, but its carbon emission cost is the highest among all forty-five SP routes. In contrast, the route selected by GA for the same transportation task, "176-126-20-284-114-102-48", although longer at 17.57 kilometers, achieves significant control over carbon emission costs by intelligently avoiding potentially severe congestion segments.

3.4. Comparative Analysis Between GA and SP Methods

To comprehensively evaluate the effectiveness of the GA optimization method, we conducted a detailed comparative analysis of the results from both GA and SP methods. Table 3 summarizes the comparative data between the two methods in terms of total cost and carbon emission cost.

Table 3 Comparisons between GA and SP methods

Method	Cost		Carbon Emission Cost	
SP	503231.53		7476.40	
GA	388549.64	-22.79%	5496.23	-26.49%

Comprehensive analysis indicates that while the traditional shortest path algorithm has advantages in geographical distance, it often leads to higher total costs and carbon emission costs in complex traffic environments due to insufficient parameter consideration. The main deficiency of the SP method lies in its failure to incorporate the additional carbon emissions caused by traffic congestion into consideration. In contrast, the GA method, by comprehensively integrating multiple factors such as distance, carbon emission costs, and transportation costs, is able to search for optimization solutions in a broader solution space, thereby achieving significant reductions in overall transportation costs.

4. Conclusion

This research presents a comprehensive methodology for addressing the Green Vehicle Routing Problem (GVRP) in urban transportation environments. The study employs a multi-stage approach combining K-means clustering for traffic pattern recognition, genetic algorithm optimization for route planning, and traditional shortest path algorithms as performance benchmarks.

The K-means clustering successfully categorized 297 Beijing traffic data points into three congestion levels, providing a realistic foundation for incorporating traffic conditions into route optimization. The spatial distribution of these clusters aligns well with actual urban traffic patterns, with low congestion areas in periphery regions, medium congestion in middle rings, and high congestion concentrated in city centers and transportation hubs.

The genetic algorithm demonstrated superior performance across all scenarios, achieving significant cost reductions compared to traditional shortest path methods. Notably, the GA approach reduced total costs by 22.79% (from 503,231.53 yuan to 388,549.64 yuan) and carbon emission costs by 26.49% (from 7,476.40 yuan to 5,496.23 yuan). These results highlight the algorithm's ability to balance multiple optimization objectives, including distance, carbon emissions, and operational costs.

The comparative analysis reveals that while shortest path algorithms excel in minimizing geographical distance, they fail to account for real-world traffic complexities, resulting in higher overall costs and environmental impacts. The GA method's strength lies in its comprehensive integration of multiple cost factors and its capability to navigate complex solution spaces to identify truly optimal routes that consider both economic and environmental sustainability. This research demonstrates that intelligent route optimization can significantly contribute to green logistics and sustainable urban transportation systems.

References

[1] Hong J, Zhan C, Lau K H. Leveraging joint distribution in urban express delivery to lessen environmental impacts: a case study[J]. Nankai Business Review International, 2022, 13(4): 567-586.

[2] Guo X, Zhang W, Liu B. Low-carbon routing for cold-chain logistics considering the time-dependent effects of traffic congestion[J]. Transportation Research Part D: Transport and Environment, 2022, 113: 103502.

[3] Chen J, Liao W, Yu C. Route optimization for cold chain logistics of front warehouses based on traffic congestion and carbon emission[J]. Computers & Industrial Engineering, 2021, 161: 107663.

[4] Zhou X, Liu C, Zhou K, He C, Huang X. Improved ant colony algorithm and modelling of time-dependent green vehicle routing problem[J]. Journal of Management Sciences in China, 2019, 22(5): 57-68.

[5] Zhao Z, Li X, Zhou X. Green vehicle routing problem optimization for multi-type vehicles considering traffic congestion areas[J]. Journal of Computer Applications, 2020, 40(3): 883-890.

[6] Xiao Y, Konak A. A genetic algorithm with exact dynamic programming for the green vehicle routing & scheduling problem[J]. Journal of Cleaner Production, 2017, 167: 1450-1463.

[7] Li K, Li D, Ma H Q. An improved discrete particle swarm optimization approach for a multi-objective optimization model of an urban logistics distribution network considering traffic congestion[J]. Advances in Production Engineering & Management, 2023, 18(2): 211-224.

[8] Zhang S, Zhou S, Luo R, Zhao R, Xiao Y, Xu Y. A low-carbon, fixed-tour scheduling problem with time windows in a time-dependent traffic environment[J]. International Journal of Production Research, 2023, 61(18): 6177-6196.

[9] Yang L, Gao Y, Sun Y, Li J. Two-Phase Hybrid Search Algorithm for Time-Dependent Cold Chain Logistics Route Considering Carbon Emission and Traffic Congestion[J]. IEEE Access, 2024, 12: 95128-95151.

[10] Li Y, Lim M K, Tan Y, Lee S Y, Tseng M L. Sharing economy to improve routing for urban logistics distribution using electric vehicles[J]. Resources, Conservation and Recycling, 2020, 153: 104585.