Delivery Centre Optimization for Last Mile Delivery

Modelling, Simulation and Optimisation (MSCDAD_B)

Debayan Biswas
Student ID: 22242821
Master of Science in Data Analytics
National College of Ireland
x22242821@student.ncirl.ie

***Abstract*—This project report gives a detailed idea about the simulation-based optimization approach for last-mile delivery service. Last-mile delivery defines the delivery of goods from a fulfillment center to the end-user or customer. This study details the problems associated with parcel distribution efficiently from warehouse to customers. The project aims to provide an optimal delivery center by running simulations of the scenario considering customer location, delivery route, shortest path for delivery, operational cost, prioritization, etc. The optimal delivery center location is achieved by running a synthetic workload simulation for a specific period and then evaluating the statistical reports to justify the decision. The simulation and optimization are achieved by running Python programming language in Jupyter Notebook.**

***Keywords*—simulation, optimization, last-mile delivery, warehouse, heuristic**

I. Introduction

The project defines the challenges the startup company named We-Doo faces in present-day logistic operations in an urban setup. Logistics operations nowadays are getting more complex and complicated as more people like to have the deliverables delivered to their homes without any hassle. This scenario is increasing day by day. Companies need to earn a maximum profit from the service they are providing to the customers by reducing operational costs, reducing the time of delivery, and enhancing the overall customer service. The most important point to consider here is to have an ideal delivery location center which might reduce the extra cost associated. Thus, the primary objective of this project is to locate an optimal delivery center location from a set of multiple locations in an area to deliver the deliverable to the customer. The optimal delivery center locations need to be located within an area to reduce the time of delivery, reduce operational costs, and enhance overall service to customers. This report proposes a simulation-based method to generate an optimized delivery center location for last-mile delivery which can be implemented in any setup in the future.

The approach of this project consists of proper consideration of various parameters such as generating a map for a location, generating customer locations on the map, creating delivery data, generating warehouse location, considering the shortest path and the shortest delivery route, parcel prioritization, considering other parameters like working time, route length, parcel leftover if any, and most importantly the optimization of delivery location. These parameters make the complete setup for generating an ideal and optimized delivery center and are generated using a seed of 2821 which is extracted from the last four digits of the student ID of 22242821.

II. Literature Review

A. *Traveling Salesman Problem*

The Traveling Salesman Problem (TSP) ensures the best route to traverse to reach all available nodes in a map with the highest efficiency and minimum operational cost. The author [1] analyzes the performance of last-mile delivery based on a combination of a truck and multiple drones using the traveling salesman problem (TSP) model. The model generated is used to minimize the total delivery time for a parcel. The delivery is based on a single truck and the use of multiple drones in combination with a two-staged heuristic formulated K-means and metaheuristic approach. Time-saving is achieved by using the model is 7.13% to 24.17% compared to truck mode. This model reduces the tour time to a great extent. However certain parameters are involved that need to be fulfilled to gain efficiency. The TSP model is applicable in areas with a high density of customers especially in urban areas. Drone endurance is also a major factor in increasing the efficiency of the model as the longer the drone speed, the faster the efficiency. The author shows that the TSP model can achieve time-saving from 3.49% to 28.20% if the drone speed changes to 80mph. This is a significant improvement.

B. *Monte Carlo Optimization*

Model optimization techniques increase the efficiency of a proposed model to implement it in real-life scenarios. The author [2] addresses the issues of reducing the damage caused by a storm to a power grid by managing efficient routing of vehicles to tackle two tasks: investigating grid-related damage and damage repair. The author uses a Monte Carlo tree search-based model which provides improved results than the standard heuristics approach. The author proposed a new method of vehicle management with respect to performing physical tasks while handling uncertain parameters at the same time using a tree search-based model. However, the model is limited to managing a single resource at a time.

C. A-star algorithm in pathfinding

The advancement of artificial intelligence can pave the way for intelligent pathfinding which can optimize work in various sectors like robot walking routes, movement of NPCs in the game, improved last-mile delivery, and so on. The author [3] studies the workings of the A-star algorithm and compares its effectiveness over the Dijkstra algorithm using the Raster method to simulate a real-world-like environment. The author uses search traversal algorithms like depth-first search (DFS) and breadth-first search (BFS) to traverse the depth and expand all nodes in a graph or a map generated. Both the deep searches offer faster time but do not ensure the shortest path always. This is taken care of by the A-star algorithm. A-star algorithm uses evaluation to predict the nodes to search and considers the nodes to be prioritized overall implementing optimized shortest path finding.

III. Methodology

The methodology for finding the optimized delivery center location consists of a step-by-step approach to simulate the daily operations of last-mile delivery services. The process included the generation of map data for a location, the generation of delivery data, finding the shortest path and delivery route, model verification, the generation of working time, route length, and parcel generation, followed by the most important step of model optimization.

A. Generating Map Data

The first step involves the generation of map data with the customer's location representing a town map. The map consists of nodes, vertices, and edges which are generated randomly based on the parameters. The graph is generated based on the random seed of 2821 which is the last four digits of the Student ID 22242821. The graph data is stored in a pickle file which can be used during later operations. This map will represent any location by just tweaking the internal parameters of the map.

B. Generating Delivery Data

A random location of the Warehouse at one of the intersections or vertices on the map is generated. The delivery data is generated based on parameters such as the average number of parcels per day for each customer, the number of customers to be served, and the number of days for which the simulation will be generated. This generated delivery data will be implemented for the model optimization in the later stages. The seed of 2821 which is the last four digits of the student ID 22242821 is also passed into the function to check delivery data for a period of 5 days and for a longer period of 30 days. The function then generates the total delivery data. The sum of the total number of parcels to be delivered is also generated by summing up the lengths of all inner listings of the delivery data.

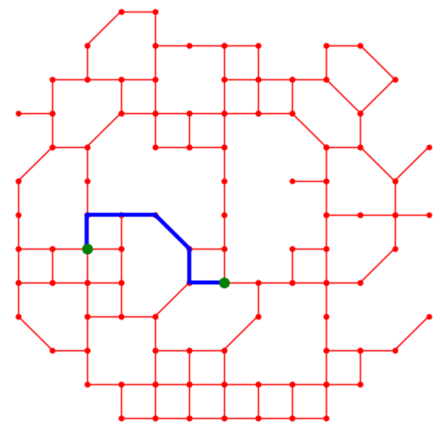


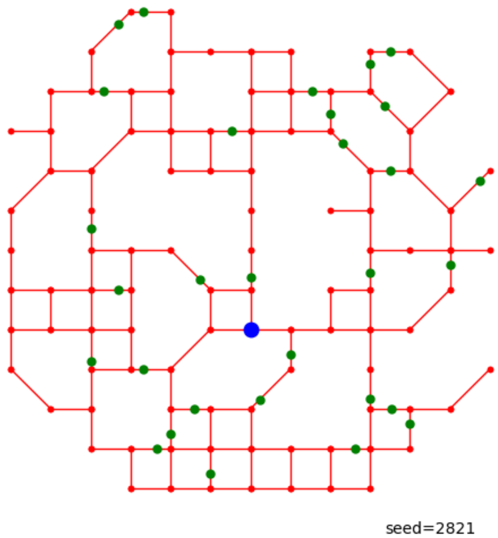
Fig. 1: Shortest Path between two random nodes

C. Finding Shortest Path

A shortest path algorithm is used to find the shortest route between two nodes in the map generated. There are many path-finding and graph traversal algorithms present such as Dijkstra and A-star algorithm. The a-star algorithm is used here as the shortest path-finding algorithm. This is because of the fact that A-star finds a better path using a heuristic approach that involves prioritizing better-performing nodes which is not the case for Dijkstra. The shortest path function calculates the shortest route between two locations or nodes on the map. The efficiency of the shortest path is evaluated to find the optimal path using the seed of 2821. Fig. 1 shows the shortest path between two customer locations generated on the map. This will be used in the process of generating the optimal location in the later stages. The shortest delivery path from the warehouse to each customer location is determined, optimizing delivery routes. The next step involves the generation of the shortest delivery path.

D. Finding Shortest Delivery Path

This step involves optimizing the delivery route by finding the shortest delivery path from the warehouse to each customer's location. A target location on the map is generated along with a warehouse location. This process uses the same A-star algorithm as a function and also uses functions such as Loops and Round trips to reach all the customers. The shortest delivery path is located from a random warehouse location that can serve all the customers. This is an optimization-based approach. A heuristic approach is also implemented to find the shortest path. The author [4] stated that an effective and efficient result can be obtained using a heuristic approach but might not always guarantee a solution that is more optimal. Their study on a combined truck-drone heterogeneous environment provided a better time-optimized solution. The heuristic approach in our project generates a path that covers a lot more distance than the optimized approach thereby increasing the cost of the operation and time. Considering this approach will not provide an optimized result and thus



(a)

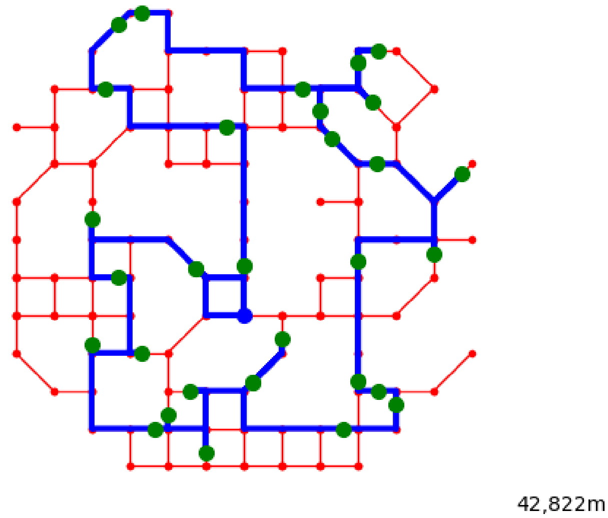
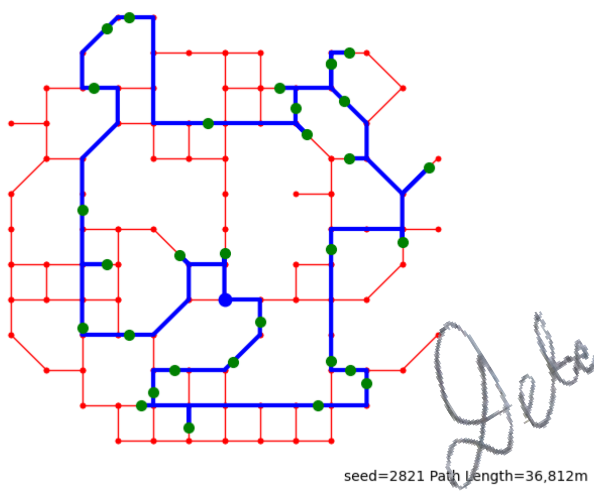


Fig. 3: Heuristic approach to find shortest delivery route



(b)

Fig. 2: Shortest Delivery Path optimization-based approach

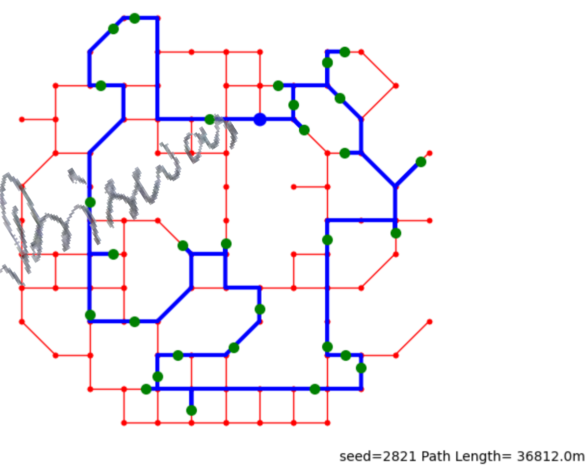


Fig. 4: Monte Carlo Optimization to find shortest delivery route

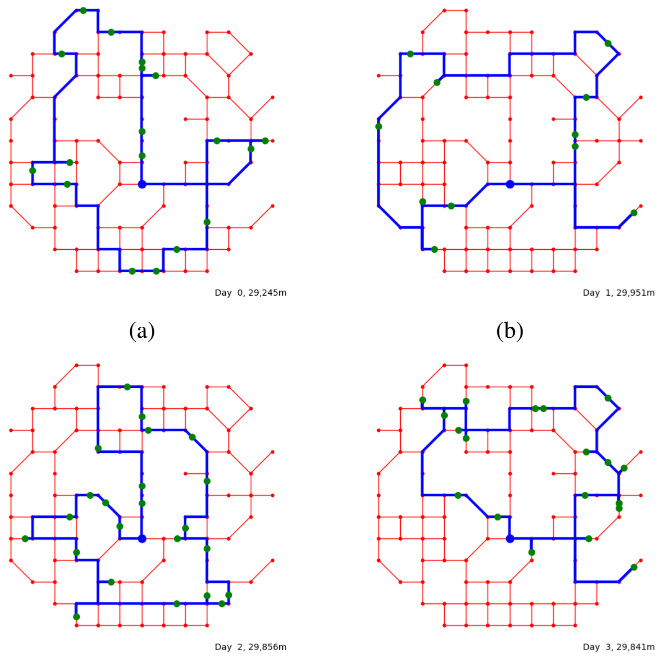
this can't be considered. This is followed by a more efficient method called the Monte Carlo Optimization method which is used to identify the optimal location of the warehouse and delivery route by creating and evaluating multiple possible solutions. The algorithms are tested with the seed of 2821 which is the last four digits of the student ID 22242821 to find its effectiveness. The Fig. 2a shows the plot which consists of a blue marker which is the warehouse location and the green markers which are the customer location. The loop function now creates a shortest delivery path simulation to depict the shortest possible routes to reach to the customer location from the warehouse which is shown in Fig. 2b.

The Heuristic approach as can be seen in Fig. 3 covers far more distance than the optimized approach thereby increasing the cost of daily operations. The increased metrics can't be associated with an optimized location so this method is not considered. Monte Carlo Optimization method creates a more optimized location by considering all possible solutions to find the shortest delivery route which can be seen in Fig. 4. These sets of algorithms act as tools for optimization and improvement of delivery efficiency.

E. Model Verification

A Verification test is performed to find and verify the accuracy of the simulation model. The simulation model is created to simulate the daily operations. Time handling function is introduced to handle time miss-match and to make sure there is no discrepancy related with time.

This step involves the addition on new classes for Parcel, Customer, Driver, and Delivery Center. These classes hold all the necessary steps to justify the process of simulation and optimization. A simulation is generated based on the input parameters and classes. The simulation with a seed of 2821



(a) (b) (c) (d)

Fig. 5: Simulation of Model

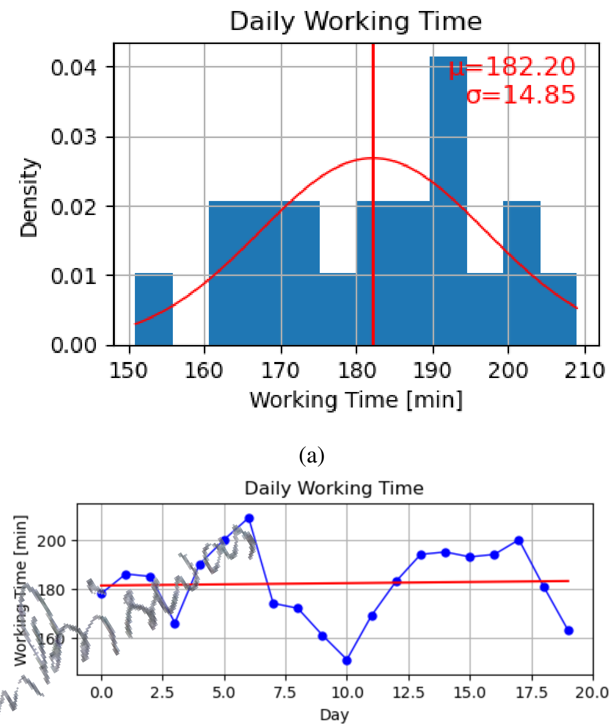
is generated for four days to verify the model's accuracy. The Fig. 5 shows the plot for 4 consecutive days starting from Day 0 to Day 3.

F. Adding Visualization and Statistics

This step deals with further data visualization and statistical analysis to interpret the results of the simulation. The distribution of the histogram and plot generated by the methods working time, route length, parcel leftover, and their respective statistics gives a proper idea about the simulation results. The histogram and plot show the distribution of working time, route length, and parcel leftover. The statistical measure of the three methods gives the mean, median, mode, and standard deviation. The distribution of the working time is a normal distribution. The Tour length generates a distribution that is left-skewed or negative distribution. This indicates that there is a tendency to have higher values which means a higher number of observations with higher value is present than that of lower value. This states that the more the tour length increases, the higher the observation. The Left-over parcel generates a right-skewed distribution which indicates the data has a low value tendency which means a higher number of observations with lower values are present.

G. Optimization

This is the final process to find the optimal delivery location using optimization techniques which is in accordance with minimizing operational costs. The Monte Carlo optimization procedure is applied to find the optimal location of the warehouse within a map. This method iterates over warehouse location by calculating the mean tour length for each of the locations and then selects the position with the least value. The seed value of 2821 which is the last four digits of the student ID 22242821 is passed to the Monte Carlo optimization function along with a p-value of 0.28 which identifies the best warehouse position out of all the warehouse positions generated.

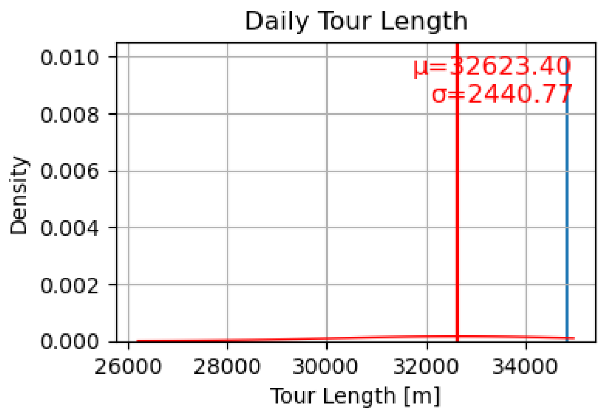


(b)

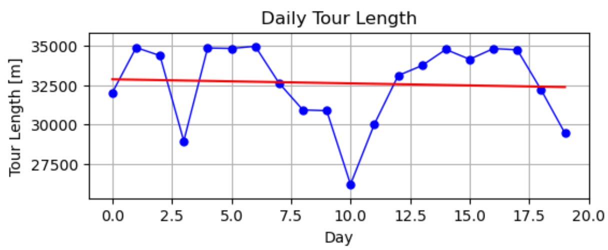
Fig. 6: Working time

IV. Results and Interpretation

The project considers the optimization-based approach rather than the heuristic approach because of the fact that heuristic one covers a greater length of distance that increases the cost of the overall operations and hence it will not impact the optimization. The result of the simulation provides an optimal location candidate warehouse location along with important statistical information regarding the simulation. The histogram in Fig. 6a shows the distribution of the daily working time of the driver for 20 days. The distribution provides an average working time of 182.20 with some variation between days which is given by a standard deviation of 14.85. The distribution of the working time depicts a normal distribution. The line plot in Fig. 6b shows the working time for each and every day over a period of 20 days. The plot clearly shows that Day 10 generates the least working time which is also proportional to the distance covered for the same location. On checking the simulation record, it is seen that on Day 10, the driver has traveled the least distance thus the working time is also less. Day 6 shows the highest working time among the other days and is above



(a)

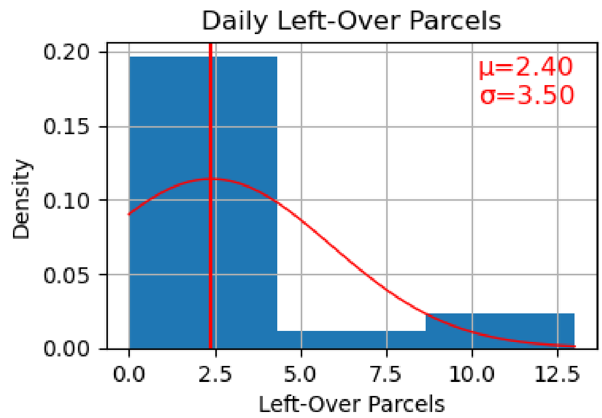


(b)

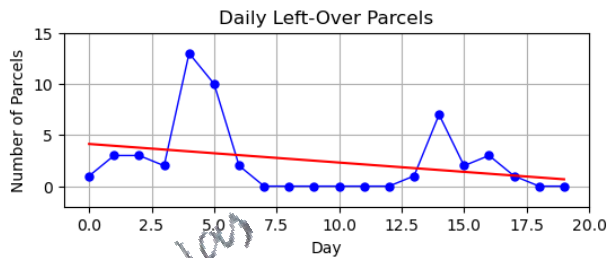
Fig. 7: Route or Tour Length

200 minutes.
The histogram in Fig. 7a shows the distribution of the Daily tour length of the drivers for 20 days. The histogram data reveals that the average distance traveled by the driver is 32623.40 meters with a variation of 2440.77 meters which is the standard deviation. The Tour length generates a distribution that is left-skewed or negative distribution. This indicates that there is a tendency to have higher values which means a higher number of observations with higher value is present than that of lower value. Fig. 7b is a line plot that shows the daily tour length. The plot shows the minimum tour length seen on Day 10 which exactly matches with the simulation value of Day 10 where the driver covered a distance of 26201 meters which is way less than the distances covered on all the other days. The highest distance covered in the span of 20 days is on Day 6 which is 34962 meters.

The histogram in Fig. 8a shows the density distribution of the daily left-over parcels. The histogram shows that the average left-over parcel is 2.40 with a variation of 3.50. The left-over parcel generates a right-skewed distribution which indicates the data has a low value tendency which means a higher number of observations with lower values are present with time. The line plot in Fig. 8b shows left-over parcel count per day. The left-over count suddenly rises and falls from Day 3 to Day 7. After that, the count of the left-over parcels follows a negative trend which in turn states that the number of left-over parcels is decreasing with time thus indicating delivery and operational efficiency.



(a)

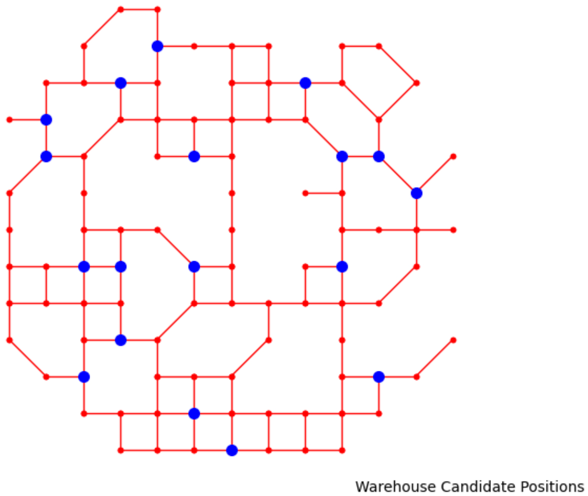


(b)

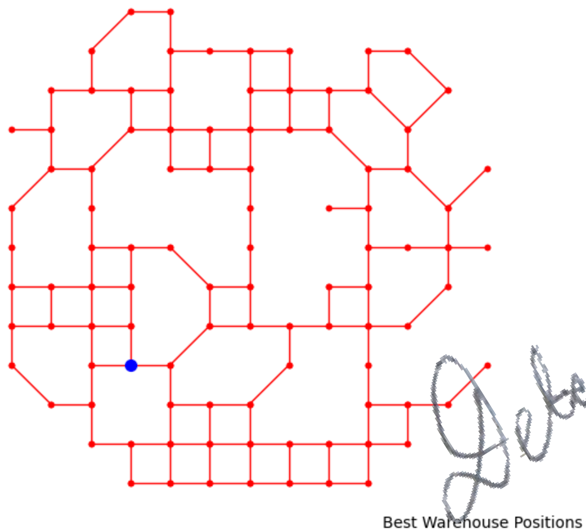
Fig. 8: Parcel leftover

The optimization is now simulated for 20 days to generate the best warehouse locations for optimized delivery to the customers. This is the terminal process to pinpoint the most optimized location out of the best locations generated on the map. The Fig. 9a shows the 18 best-candidate warehouse locations that are generated. The simulation runs Monte Carlo optimization for each and every candidate warehouse location generated for 20 days to finally point the best delivery location. The best location is found which can be seen in Fig. 9b with a location coordinate of (2320, 2320). This is the most improved location co-ordinate out of the total 18 locations which gives the most optimized location for delivery with a mean cost of 94.18706.
The next step is to compare 5 random locations from the 18 optimized locations generated and run a statistical test to compare and contrast the locations and determine the best location. Fig. 10 shows the list of mean cost and mean work time of 5 random locations. Among these, the warehouse location with the least mean cost can be considered the most optimized, but the difference is still tiny compared to the other locations.

Statistical analysis is implemented to further justify the results. A combined Anova test for cost function and working time is calculated for all the locations which can be seen in Fig. 11. The test provided a p-value of 0.99984 which is very close to zero and an F-statistic value of 0.0087. The



(a)



(b)

Fig. 9: Final Optimized result

p-value suggests there is no significant difference between the groups of locations that are being compared and the F-stat value indicates no significant difference between the groups. Thus the optimized location obtained from Monte Carlo optimization can be concluded as the best optimal location.

V. Reflections and Future Work

There is a wide array of scope present that can be implemented over this project to enhance its operational capability. Additional factors that can influence the optimum delivery can also be considered such as traffic congestion, road conditions, changes in real-time demand, demography and topography of the location, and so on. The scope lies in the improvement of the simulation model by adding real-time data to increase the accuracy of the optimization strategy.

```
Mean Daily Costs:
Location: (1760, 1760), Mean Daily Cost: 94.14868000000001
Location: (5680, 5120), Mean Daily Cost: 94.331724
Location: (2320, 6240), Mean Daily Cost: 94.213632
Location: (2320, 2320), Mean Daily Cost: 94.18706
Location: (2320, 3440), Mean Daily Cost: 94.52292399999997

Mean Working Times:
Location: (1760, 1760), Mean Working Time: 183.05
Location: (5680, 5120), Mean Working Time: 183.4
Location: (2320, 6240), Mean Working Time: 183.2
Location: (2320, 2320), Mean Working Time: 183.15
Location: (2320, 3440), Mean Working Time: 183.8
```

Fig. 10: Mean cost and mean working time for 5 random locations

```
ANOVA for Daily Costs:
F-statistic: 0.008705667696906595
P-value: 0.9998470907928391

ANOVA for Working Times:
F-statistic: 0.008781232488596313
P-value: 0.9998444410872506
```

Fig. 11: Statistical test of 5 random locations

The model can further be enhanced by using machine learning algorithms that will learn from past data and will implement future actions accordingly, leading to more efficient deliveries. There is also a scope to implement sustainability in the future by adapting to eco-friendly techniques to optimize delivery by maximizing fuel efficiency and reducing carbon emissions of delivery vehicles by using more efficient nature-friendly transportation.

Making the model more flexible to changes can make it more adaptable for future operations and will pave the path for the overall success of the simulation model.

Acknowledgments

The support, guidance, and critical feedback by Professor Hicham Rifai led to the completion of the project.

References

[1] S.-H. Lu, R. Kuo, Y.-T. Ho, and A.-T. Nguyen, "Improving the efficiency of last-mile delivery with the flexible drones traveling salesman problem," *Expert Systems with Applications*, vol. 209, p. 118351, 2022.

[2] L. Al-Kanj, W. B. Powell, and B. Bouzaiene-Ayari, "The information-collecting vehicle routing problem: Stochastic optimization for emergency storm response," 2023.

[3] Y. Yan, "Research on the a star algorithm for finding shortest path," *Highlights in Science, Engineering and Technology*, vol. 46, p. 154–161, Apr. 2023.

[4] M. Rinaldi, S. Primatesta, M. Bugaj, J. Rostáš, and G. Guglieri, "Development of heuristic approaches for last-mile delivery tsp with a truck and multiple drones," *Drones*, vol. 7, no. 7, 2023.