Airport Cab Drivers' Passenger Carrying Strategy Selection Based on Mathematically Derived Theoretical Models

Yiqi Lu*, Jiahao Wang

School of Mechanical and Automotive Engineering, Guangxi University of Science and Technology, Liuzhou,545616, China

*Corresponding author: 100002187@gxust.edu.cn

Abstract. Cab is an important way for passengers to choose to go to the city in the airport. Cab drivers will send customers to the airport. There are two options to carry passengers, choosing to return to the city to carry passengers or entering the pool at the airport to wait for passengers. How the cab driver to decide to rely on the pool waiting time, the airport of a single day by the time of day the flow of passengers, downtown to the airport in the time required, and other factors. To compare the two options and analyze the influencing factors, this paper establishes a theoretical model of mathematical derivation to find out the scope of application of the two types of options, and to establish a decision-making model for cab drivers to choose. Then, using specific airport data, MATLAB is used to fit the curve to the collected data, and the decision value is calculated by applying integrative science to get the cab's time from the city to the airport, to arrive at the decision-making scheme. Finally, the dependency of the relevant factors is analyzed and the effect of different times of day, season, and fuel cost government-related policies on the revenue coefficient is considered using gray correlation analysis.

Keywords: Gray Correlation Statistics, Mathematical Derivation Theory, Loading Strategy.

1. Introduction

With the acceleration of urbanization, airport construction has become an important part. The rapid development of the air transportation industry has brought about a rapid increase in passenger travel, and constantly put forward higher requirements for the configuration of the matching transportation system[1]. Cab drivers stationed at the airport have two options for carrying passengers to choose from, returning directly to the city or entering the queue of the airport storage pool, the main factor affecting the decision of cab drivers is the revenue per unit of value, and the factors that have a greater impact on the revenue are the different periods, the number of landings and takeoffs at the airport, seasonal factors, important meetings, and weather factors. Among the known analysis methods, the integer linear programming model is used, which makes the required vehicles as few as possible, the waiting time of passengers as short as possible, and the mileage of passenger ride detours as few as possible[2]. Some scholars construct a decision support model from the perspective of opportunity cost, taking into account the relevant factors affecting the cab driver's decision-making, and constructing a linear logarithmic regression model of the passenger's willingness to take a cab and the cab driver's opportunity cost[3]. Based on the analysis of the factors affecting drivers' decision-making, the synergistic association between the number of flight arrivals in different periods and the average cab-seeking distance and other core factors is deeply explored[4]. Based on the cost-benefit analysis method and multilevel grid division algorithm, the factors affecting the decision-making of cab drivers were identified [5,6]. The mathematical methods such as functional equations, BP neural network algorithms and queuing theory were used to establish a complete decision-making model for cab drivers and an optimization model for the efficiency of airport scheduling [7,8]. In the process of research and analysis of the problem, the basic methods of information theory were browsed to determine the basic ideas on the selection of influencing factors and decision-making approaches [9].In terms of information organization, curve fitting of discrete information can be used to find out the correlation curve for further research. In the field of decision-making research, decision-making research in other industries is also of great significance. Focusing on the study of air quality influencing factors in the field of environmental science, the gray correlation between motor vehicle ownership and urban air quality concentration of each monitored pollutant was analyzed using gray absolute correlation analysis, gray relative correlation analysis and other methods [10]. In the study of operation and management of rail transit, multi-attribute decision-making theory was selected as the main idea of problem solving, and the decision-making attributes were determined through literature research and analysis [11].In the optimization problem of energy saving and emission reduction in the field of construction, the selection problem of multiple solutions also determines the advantages and disadvantages of building design and structural design. Some scholars in this field of research have made a comprehensive and detailed compendium of optimization methods and optimization tools based on agent models, and introduced a multi-attribute decision-making method based on subjective-objective combination of empowerment-TOPSIS evaluation for evaluating and selecting building performance optimization schemes [12]. The research on decision-making methods in various fields also provides new ideas and data processing methods for this study.After considering the above factors and scholars' studies, the relationship between the total revenue coefficient and the time for cabs to travel from the city to the airport as well as the time for cabs to wait in the storage pool was quantitatively analyzed, and the gray correlation method was applied to analyze the effects of different periods, seasons, fuel costs, and relevant government policies on the revenue coefficient every day.

2. Modeling and Calculation of Option Selection

2.1. Establishment and calculation of the coefficient of return for the two programs

After picking up a passenger from the city and delivering it to the airport, the cab driver chooses to wait in line at the storage pool to pick up the passenger before returning to the city. This will involve three processes, which are shown in Figure 1.

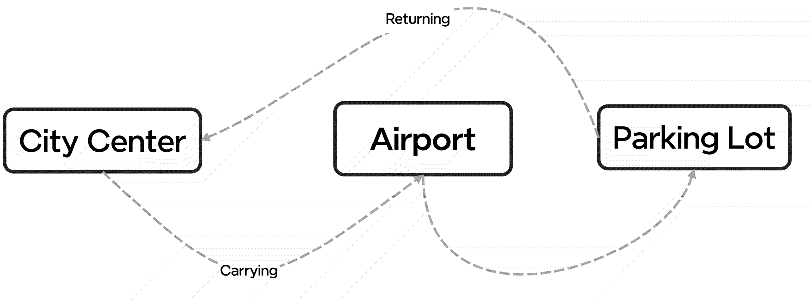

Figure 1. Queuing process for the vehicle storage pool

(1) The process of driving from the city to the airport;
(2) The process of waiting at the storage pool to pick up passengers;
(3) Returning to the city after picking up passengers.
A round trip by cab from the city to the airport yields.

$$W_A = a_1x_1 + a_2x_2 + a_3x_3 - c_A \tag{1}$$

The total return factor for Option

$$\lambda_A = \frac{a_1x_1 + a_2x_2 + a_3x_3 - c_A}{x_1 + x_2 + x_3} \tag{2}$$

a_1 : Downtown to Airport Driver Revenue Factor

a_2 : Queuing Passenger Benefit Factor for Storage Pools

a_3 : Return - to - urban yield factor for passengers

c_A : Requirements for road and fuel costs

x_1 : Time from city to airport

x_2 : Time required to wait in line for passengers at the airport

x_3 : Time required to return to the city after loading passengers

A cab driver picks up passengers from the city delivers them to the airport, and immediately returns to the city empty to pick up passengers. The system involves two processes, which are represented in Figure 2.

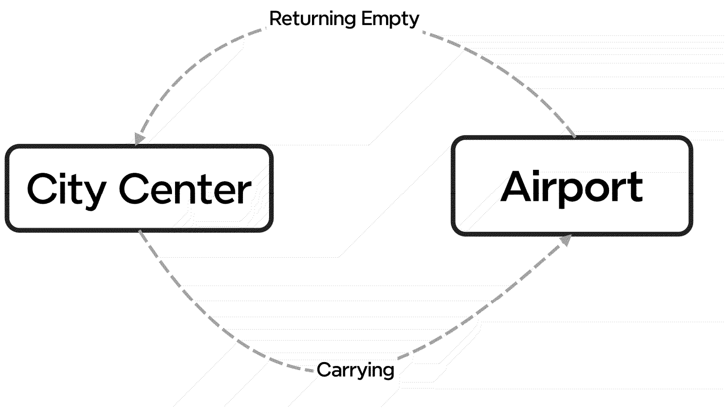

Figure 2. Drop off passengers and return to the city by empty taxi

(1) The process of driving from the city to the airport;

(2) The process of returning empty from the airport to the city after dropping off a passenger.

A cab makes one round trip from the city to the airport.

$$W_B = a_1x_1 + a_4x_4 - c_B \tag{3}$$

Adoption of total return factor for option

$$\lambda_B = \frac{a_1x_1 + a_4x_4 - c_B}{x_1 + x_4} \tag{4}$$

a_1 : Cab yield factor from city to airport

a_4 : Return - to - urban yield factor for empty vehicles

c_B : Road and fuel costs

x_1 : Time from city to airport

x_4 : Time required to return empty

According to the survey data, the starting price of a cab is about 8 yuan; In addition, cabs charge another 2 yuan per kilometer; The speed of cabs on the road is about 50 kilometers per hour, and the cost of fuel is 0.3 yuan per kilometer.Cab fare and cost information is shown in Table 1.

Table 1. Taxi fare information sheet for a second-tier city

Single Mileage Earnings(Data from the forecast of a second-tier city in China)	
Starting price	8 yuan
Rental price per kilometer	2 yuan per kilometer
Starting price kilometers	3 kilometers
Fuel costs	0.3 yuan per kilometer
Highway toll costs	0.4 yuan per kilometer
Average travel speed	50 kilometers per hour

Considering only the fuel cost and not other fixed costs, the following coefficients of return are derived by applying the principle of integration based on the above formula and the analysis of a large amount of data:

$$a_1 = \text{Starting price} + 2\times 50 \text{ - fuel cost}$$

$$a_2 = \text{Average earnings per hour} = \text{-}19$$

$$a_3 = a_1 = 93$$

$$a_4 = \text{fuel cost} = \text{-}15$$

Substitute each gain coefficient into Equation (1) (2) (3) (4):

$$\begin{cases} W_A = 93x_1 - 19x_2 + 93x_3 - c_A \\ W_B = 93x_1 - 15x_4 - c_B \end{cases} \tag{5}$$

$$\begin{cases} \lambda_A = \dfrac{93x_1 - 19x_2 + 93x_3 - c_A}{x_1 + x_2 + x_3} \\ \lambda_B = \dfrac{93x_1 - 15x_4 - c_B}{x_1 + x_4} \end{cases} \tag{6}$$

2.2. Programmatic decision-making

Assuming zero fixed costs such as fuel and road costs, and equal round-trip times, the size of the total return coefficients are compared as a means of determining the driver's chosen decision option.

When $\lambda_A > \lambda_B$, solve for $x_1 : x_2 > \frac{186}{131}$ and choose the carpool option.

When $\lambda_A < \lambda_B$, solve for $x_1 : x_2 < \frac{186}{131}$ and choose the return to downtown option.

When $\lambda_A = \lambda_B$, solve for $x_1 : x_2 = \frac{186}{131}$ and choose both options.

Cab drivers' earnings and downtown-to-airport times are related to waiting times at the storage pool.

3. Selection Decision Modeling and Solving

3.1. Modeling Using Curve Fitting

The typical travel time from the city to the airport is 1.1 hours, and the decision-making methodology described above indicates that the maximum time required to wait in line for passengers at the airport would be 1.35 hours, which is the value used to express the "option cut-off" below.

According to the data of the waiting time of 24 small cabs at the airport in the pool, MATLAB was used to establish a curve-fitting model to react to the pattern of the waiting time in the pool during the day. The program selection demarcation line represents the demarcation line between "return to downtown" and "enter the pool". If the waiting time is above the program selection demarcation line, choosing "pool". Otherwise, the "return to downtown" option is chosen. The relationship between the length of waiting time in the storage pool and the period of a single day is shown in Figure 3.

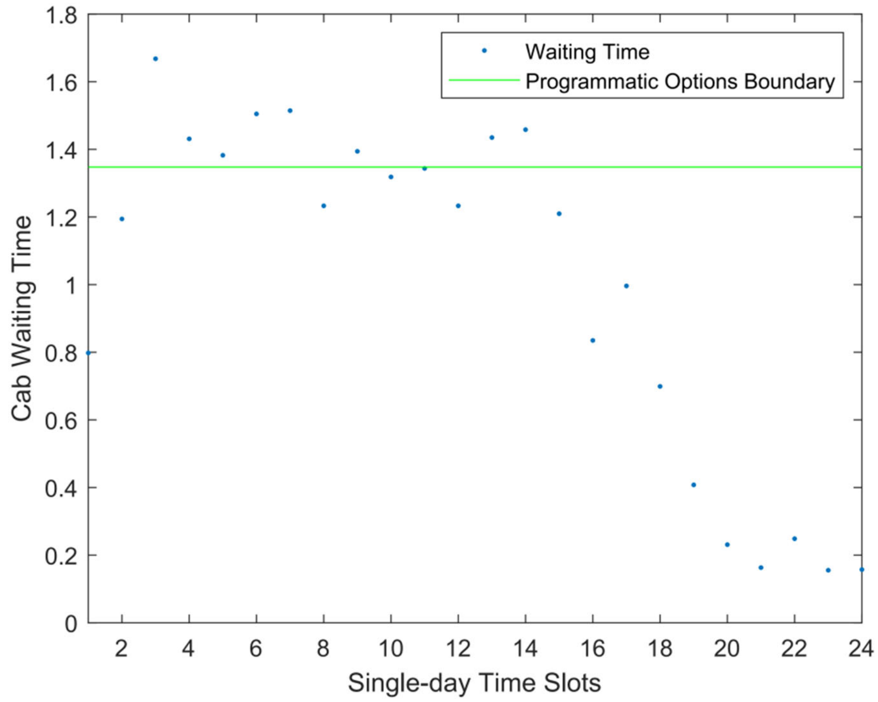

Figure 3. Plot of parking pool wait time versus single day time period

The MATLAB Basic Fitting tool is used to fit the scatter plot, which in turn leads to the equation of the fitted curve.

$$f(x) = -0.002975x^6 + 0.25441x^5 - 8.3094x^4 + 130.05x^3 - 1013.8x^2 + 3622.8x - 759.64 \quad (7)$$

The curve fitting results are shown in Figure 4.

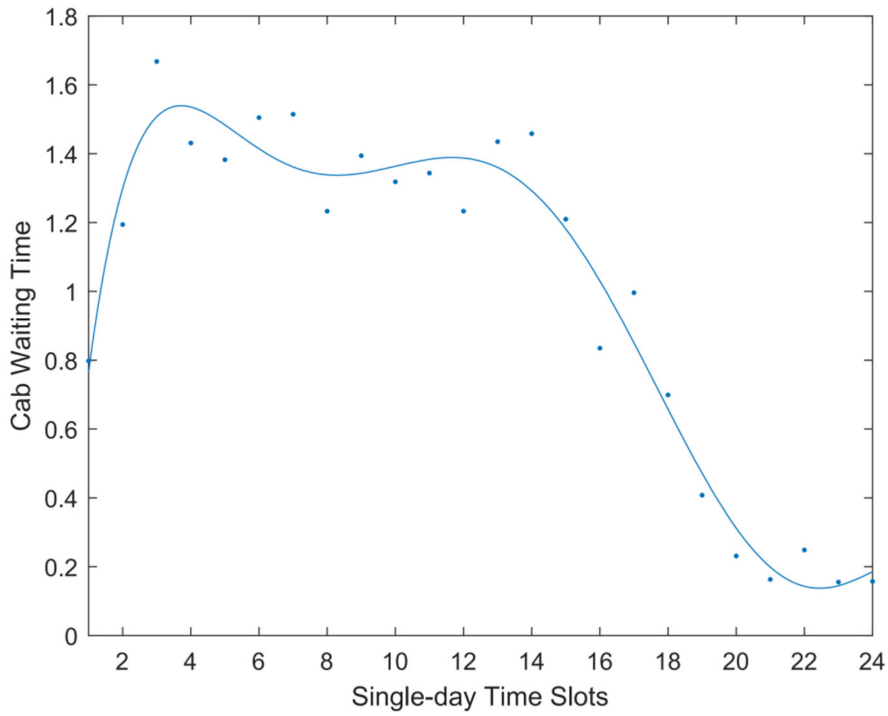

Figure 4. Curve fitting results

Discuss the relationship between the curve equation and the time required to wait in line for passengers at the airport based on the curve equation and the scenario divider. Solve for the driver's decision-making method by finding out which option the driver chooses to maximize revenue.

When $f(x) \geq 1.35$ the cab driver maximizes his return by choosing the "pool" option.

When $f(x) \leq 1.35$ the cab driver maximizes his return by choosing the "return to the city" option.

3.2. Reasonableness testing

Based on the available airport cab waiting time data, the statistics are shown in Table 2.

Table 2. Waiting time statistics by period

Waiting Hours Per Hour Throughout The Day (data based on waiting market estimates for the existing storage pool at Taiyuan Wuxu International Airport)					
1:00 to 2:00	0.23h	9:00 to 10:00	1.43h	**17:00 to 18:00**	1.23h
2:00 to 3:00	0.16h	10:00 to 11:00	1.38h	**18:00 to 19:00**	1.44h
3:00 to 4:00	0.25h	11:00 to 12:00	1.50h	19:00 to 20:00	1.46h
4:00 to 5:00	0.11h	12:00 to 13:00	1.51h	20:00 to 21:00	1.21h
5:00 to 6:00	0.16h	13:00 to 14:00	1.23h	21:00 to 22:00	0.84h
6:00 to 7:00	0.80h	14:00 to 15:00	1.39h	22:00 to 23:00	1.00h
7:00 to 8:00	1.19h	15:00 to 16:00	1.32h	23:00 to 24:00	0.70h
8:00 to9:00	1.67h	16:00 to 17:00	1.34h	24:00 to 1:00	0.41h

According to the data obtained from the above survey, it is substituted into the process of the above model, to verify the origin of the fitted curve equation. After the verification of data substitution, it can be shown that under certain conditions, the above model is reliable, reasonable, and can be used to help drivers make decisions.

3.3. Dependence of relevant factors

Suppose there are multiple comparison series.

$$X_i = \{X_i(k) \mid k = 1,2,\cdots,n\} = (X_i(1), X_i(2),\cdots,X_i(n)) \quad i = 1,2,\cdots,m \tag{8}$$

$$\xi_i(k) = \frac{\min_i \min_k |X_0(k) - X_i(k)| + \rho \max_i \max_k |X_0(k) - X_i(k)|}{|X_0(k) - X_i(k)| + \rho \max_i \max_k |X_0(k) - X_i(k)|} \tag{9}$$

$\xi_i(k)$ is expressed as the correlation coefficient of the comparison series to the reference series at a given moment in time. $\rho \in [0, +\infty)$ denoted as the resolution factor.

$\min_i \min_k |X_0(k) - X_i(k)|$ 、$\max_i \max_k |X_0(k) - X_i(k)|$ are denoted as two-level minimum difference and two-level maximum difference, respectively.

$$r_i = \frac{1}{n}\sum_{i=1}^{n} \xi_i(k) \tag{10}$$

The formula is expressed as the correlation between the series and the reference series.

The raw data were processed and represented in Table 3.

Table 3. Data processing results after the introduction of correlation

	a_1	a_2	a_3	a_4
Time Interval	8	9	8	7
Vacations	8	8	6	9
Road and Fuel Costs	7	8	7	5
Relevant Policies	9	7	9	6

Determine the reference series: $\{X_0\} = \{9, 9, 9, 9\}$

After calculating the difference between the two influencing factors is represented in Table 4.

Table 4. Data processing results after the introduction of interpolation calculations

	a_1	a_2	a_3	a_4
Time Interval	1	0	1	2
Vacations	1	3	3	0
Road and Fuel Costs	2	1	2	4
Relevant Policies	3	1	1	1

Finding the maximum value and taking the resolution factor as 0.5 for the calculation, the results are as follows:

$$\min_{i=1}^{n}\min_{k=1}^{m} |X_0(k)-X_i(k)| = \min(0,0,1,1) = 0 \tag{11}$$

$$\max_{i=1}^{n}\max_{k=1}^{m} |X_0(k)-X_i(k)| = \max(2,3,4,3) = 4 \tag{12}$$

$$\xi_2(1) = \frac{2}{1+2} = \frac{1}{3}, \quad \xi_2(2) = \frac{2}{3+2} = \frac{2}{5}, \quad \xi_2(3) = \frac{2}{5}, \quad \xi_2(4) = \frac{2}{0+2} = 1 \tag{13}$$

Introducing the results of each of the above numerical calculations, the results of each influencing factor are shown in Table 5.

Table 5. Calculation results of each influencing factor after the introduction of coefficients

	a_1	a_2	a_3	a_4
Time Interval	2/3	1	2/3	1
Vacations	1/3	2/5	2/5	1
Road and Fuel Costs	1/2	2/3	1/2	1/3
Relevant Policies	2/5	1/3	1/3	1/3

Calculate the mean value of the correlation coefficients (correlation order) for each indicator in each condition.

$$r_1 = \frac{17}{24} \quad , \; r_2 = \frac{32}{60} \quad , \; r_3 = \frac{5}{12} \quad , \; r_4 = \frac{7}{20} \tag{14}$$

From the above calculations, it is clear that the revenue coefficient is most dependent on the period, followed by vacations, road tolls, fuel costs, and relevant policies.

4. Conclusions

How cab drivers make their decisions depends on factors such as the waiting time at the storage pool. The passenger flow at the airport in a single day at each period, and the time needed to get from the city to the airport. This paper establishes a theoretical model for mathematical derivation, finds out the scope of application of each of the two types of programs, determines the decision-making boundary of the program and establishes a decision-making model for cab drivers to choose. Using the curve fitting toolbox, the decision value is calculated by applying integrative science to get the cab's time from the city to the airport, to arrive at a specific decision-making scheme. Finally, the dependence of the relevant factors of the driver's decision is discussed, using gray correlation analysis to consider the effects of different periods of the day, seasons, fuel costs, and relevant policies on the revenue coefficients, which is shown by the data analysis: the revenue coefficients are most dependent on the periods.

References

[1] Huang Depeng. Research on the optimization of cab departure traffic allocation in airport hub[D]. Beijing Jiaotong University,2022.

[2] LI Huichun,LI Zhemin,MAO Ziyang. A carpool cab dispatch model based on integer linear programming[J]. Transportation Research,2018,4(05):35-42.

[3] LIN Yao, WANG Lin, XU Zonghuang et al. A decision support model for airport cab "waiting-returning" - A case study of Nanjing Lukou International Airport[J]. Computer Application and Software,2023,40(08):24-31+79.

[4] Xiaobing Ding, Zhigang Liu, Gan Shi, Hua Hu, Jiaping Chen, Kaihe Yang, Su Wan, Jinlong Wu, "The Optimization of Airport Management Based on Collaborative Optimization of Flights and Taxis", Discrete Dynamics in Nature and Society, vol. 2022, Article ID 4542299, 16 pages, 2022.

[5] Yingmiao Qian, Shuhang Chen, Jianchang Li, Qinxin Ren, Jinfu Zhu, Ruijia Yuan, Hao Su, "A Decision-Making Model Using Machine Learning for Improving Dispatching Efficiency in Chengdu Shuangliu Airport", Complexity, vol. 2020, Article ID 6626937, 16 pages, 2020.

[6] XU Fangfang,DAI Dameng. Research and implementation of global scheduling strategy for cabs based on multilevel grid division algorithm[J]. Computer Application Research,2021,38(01):115-119.

[7] Qiaoli Yang, Bo Yang, Zheng Qiao, Min-an Tang, Fengyang Gao,Impact of possible random factors on queue behaviors of passengers and taxis at taxi stand of transport hubs,Physica A: Statistical Mechanics and its Applications,Volume 580,2021,126131,ISSN 0378-4371.

[8] Shi Kewei, Zhang Xinyue, Chen Runhuan. Queueing theory model-based cab scheduling problem for car parks[J]. Electronic Technology and Software Engineering,2020(02):198-200.

[9] Bao-Yu Tian,Jie Yang,Zhi-Qiang He et al. Fundamentals of information theory[M]. People's Posts and Telecommunications Press:, 201601.314.

[10] ZHANG Wei, YANG Weisen, BAI Qian et al. A study on the correlation between motor vehicle ownership and urban air quality in Xi'an based on descriptive statistics and gray correlation model[J]. Transportation Research,2022,8(03):111-119.

[11] Fu Y. Research on the application decision of fully automatic operation technology in urban rail transit engineering [D]. Shandong University of Architecture,2023.

[12] Huo R. Research on multi-objective optimization decision-making method for architectural design under architect's perspective [D]. Shandong University of Architecture,2023.