RESEARCH PAPER

Automobile Workshop Queue System Optimization Using Response Surface

Welly Sugianto[1]*, Reazul Haq Abdul Haq[2] & Mohd Nasrull bin Abdol Rahman[3]

ABSTRACT
The automobile workshop queue system has been optimized using various approaches, such as queuing theory, simulation, and probability. The utilization of response surface methodology (RSM) for optimizing automobile workshop queue systems is not yet established. The utilization of RSM with direct observation enables the detection of patterns of correlations between variables and responses, which are then represented through mathematical equations. The optimization process involves numerous factors that impact queue performance, which can be categorized into two parts. The number of servers, number of phases, number of workers, worker experience, and layout are classified in inner design. This study examines the relationship between two components of the outer design, specifically the arrival rate and the interarrival time. The responses analyzed are queue cost, service time, average customer waiting time, and number of customers. The findings indicate that queue costs are not reliable for establishing the optimum value due to the significant impact of the cost structure on the structure of the optimal location. This study discovered that the number of leaving customers is related to queue costs and is relevant in selecting the optimal point. This study also formulates mathematical equations for predicting the optimal point. This study emphasizes the necessity for further investigation to uncover alternative mathematical equations that can precisely predict the optimal conditions for various types of services.

KEYWORDS: *Automobile; Queue; Response surface; Optimization.*

1. Introduction

As technology advances, automobile services must also adapt to suit technological developments. An optimal workshop queue design should prioritize both efficiency and profitability. There are various steps involved in car maintenance, such as component inspections, modifications, replacements, and quality control. The mechanic will promptly address the automobiles that arrive at the facility. The mechanic inspects the components to determine the components that require repair or replacement. Subsequently, the mechanic fills up the spare parts request form and forwards it to the authorized spare parts department. The mechanic affixes the replacement component into the automobile. The mechanic conducts a comprehensive inspection of the automobile to ascertain its best operational condition [1].

The implementation of a queue system is a widely used practice across numerous services. Upon entering the service facility, the customer initiates the commencement of the queue. A queuing system is a procedure or method used to organize and manage a queue of people or entities waiting for service or access to a particular facility, service, or process [2]. However, there are some queue characteristics analyzed, such as the average waiting time for customers in the queuing system, the average queue length, the average number of customers being served in a certain period, the probability that the customer does not want to wait, the probability that there are no customers in the system, and the probability that a

---

* Corresponding author: Welly Sugianto
gd200052@student.uthm.edu.my

1. *Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, 86400.*

2. *Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, 86400.*
3. *Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, 86400.*

certain number of customers are in the system. The aim is to get an overview of the queuing system [3].

There are several main characteristics that need to be considered to build an automobiile workshop queue system, which include:

1. The process of customer arrival: Customers may arrive either singly or in groups. Customers exhibit a lack of tolerance and will abandon the queue system when the waiting duration exceeds their threshold.
2. Service time: The service time can be categorized as either constant or fluctuating. The service time refers to the mean duration of service for each customer, excluding the overall time spent by the customer in the queue system.
3. Number of servers: Servers can be single or multiple.
4. System capacity (waiting room): Restrictions can occur in the waiting room where the number of customers in the system is limited. An unrestricted queue system is characterized by the absence of a waiting room.
5. The whole population refers to the number of customers who enter the queuing system.
6. Queue orders: According to the firm policy, customers are served either individually or in batches.
7. Priority allocation: Customers, particularly existing members or those in a rush, are assigned various levels of importance.
8. Service discipline can be in the form of first come first served or last come, first served.

These factors impact several queue characteristics. For example, using too few servers at a high arrival rate will increase the number of queues in the system. Improper configurations will decrease queue efficiency, thereby affecting customer satisfaction. Implementing priorities directly affects customer satisfaction. Engaging in customer prioritization might result in a decline in overall customer satisfaction, as it may necessitate compromising the needs of other customers. Excessive service time will result in an increase in both the number of customers waiting in line and the number of customers who abandon the queue system [4].

The occurrence of queues at automobile repair shops has a detrimental impact on customer satisfaction. Excessive queue length will result in customer dissatisfaction. In addition, queuing incurs costs, such as customer waiting costs and capacity costs. The customer waiting costs are associated with the costs incurred to accommodate the queue, which may include providing space, queuing facilities, etc. Capacity costs are related to the provision of a queuing system and include personnel, equipment, resource costs, and other things. As the number of waits grows, so do the queuing costs [5].

Given this context, it is crucial to enhance the efficiency of the automobile workshop queue system to enhance customer satisfaction and minimize costs.

2. literature Review
2.1 Queue optimization

Queue theory is typically used to conduct the queue analysis. The prediction of queue performances is based on the arrival rate and service rate. These performances encompass the utilization rate, the probability of having no customers, the probability of having a specific number of customers, the average number of customers in the system [6], system utility, and average customer waiting time [1].

Optimizing queueing system performance necessitates the use of both servers and workers. Many assumptions simplify calculations [7]. There are several factors to consider, such as queue discipline, arrival rate, and service rate [6], [8], [9]. The optimization of queuing system performance includes the number of customers, and the average customer waiting time, and queue cost [10], [11].

The queue system performance optimization is achieved through the utilization of modeling and simulation techniques. To optimizing queues, researchers employ various simulation methods. Queuing systems are simulated using discrete-event simulation (DES). DES is not a perfect simulation of a queue system, but it comes close. The objectives are to rationalize the queueing process and establish the best configuration [12]. The simulation is conducted by integrating the arrival rate and service time to optimize the number of servers [13]. Strategically aiming to enhance the efficiency of the queue management system, reducing the average customer waiting time is a crucial endeavor. Different layouts are simulated to determine the optimal configuration. The selected configuration enhances both the utilization rate and the average customer waiting time [14]. Modeling the queue system uses continuous simulation. The result reveals that continuity-related mistakes decrease with system scalability [15]. The arrival rate that exceeds the service time results in long lineups of cars. The queuing model is simulated by varying the duration of service time. The service time duration is adjusted in line with the arrival of cars [16]. Complex queuing systems are typically modeled via simulation while considering a variety of

elements to improve queue performance [17]. The queue system simulation encompasses a wide range of aims, which include the prediction of consumer numbers through the utilization of fluid deterministic models [18], prediction of interarrival [19], reducing the number of customers on the line using monte carlo simulation [20], reducing number of leaving customers using multi-priority simulation [21], reducing service time using numerical analysis [22], prediction of queue characteristics using metamodels made up of linear models, random forests, and neural networks [23]. The entirety of the research incorporates two key factors, specifically the service rate and the arrival rate. Managing the queuing system poses a significant challenge due to the extensive range and quantity of cars, as well as the wide variety of maintenance duties. These obstacles are also caused by the intricacy of the system's connections between various aspects. Modifications to a particular component inside the system will inevitably exert influence on the system's overall functionality. Simulation examines intricate interconnections among many components. Simulation does not provide an optimal response but provides a comprehensive representation of the current condition of the system [24].

A comprehensive financial study is conducted to determine the breakeven threshold for each queuing system, ensuring that the most recent queuing system exhibits superior performance [25]. Additionally, a multi-priority technique is utilized to improve queuing efficiency and reduce queue costs [26]. Queues result in substantial costs, as they encompass both customer queuing costs and capacity costs. The customer queuing cost encompasses costs related to allocating space for customers who are waiting, the cost of lost productivity, and the opportunity cost. The cost of lost productivity incurred by customers waiting in line as they do not engage in productive activity. The concept of opportunity cost pertains to the foregone potential earnings resulting from the departure of customers from a queue [27]. The optimization of queuing systems involves the minimization of costs associated with queues. A sequence of simulation experiments is employed to optimize the number of servers, with the objective of reducing queue costs [28].

The optimization process also incorporates a lean strategy, which involves many techniques such as value stream mapping [29], layout design development [30], critical path approach [31], design selection [32], and queue procedure analysis [33].

Researchers also employ various operations research methodologies to optimize queue performance. Markovian queuing models are employed for the purpose of evaluating the performance of queuing systems through the examination of the interarrival time [34]. Using the birth and death procedure, the number of customers is decreased by considering the arrival rate and service rate factors [35]. The reduction in waiting time is achieved through the consideration of the arrival rate factor using a non-convex nonlinear algorithm [36]. Additionally, the queue system utilizes a non-linear mathematical modeling approach to reduce the number of customers. The determining factor is the number of servers. Expanding the server capacity results in a reduction in customer volume and waiting duration [37]. The queue system optimization also considers layout. An optimal arrangement reduces the customer waiting times and improves the efficiency of movement between sections [38]. The analysis of queue performance also incorporates several unconventional techniques, including the utilization of fuzzy environments [39], Critical path method [31], and real-time system [40]. Additionally, a linear programming approach is employed. The objective is to reduce service and waiting times. Constraint functions include batch size, and transfer volume [41]. The integration of linear programming with simulation techniques is employed to minimize both the average customer waiting time and the number of customers [42]. Dynamic programming techniques [43], learning agent-based design [44], fractional programming [45], and linear infinite model [46] are employed for the purpose of optimizing the number of servers in relation to each arrival rate. Heuristic procedures are employed to reduce queue costs, minimize completion time, and lower the average waiting time for customers [47]. The queue optimization review describes the criteria used to examine queuing systems, including servers, workers, arrival rate, interarrival time, and layout. The performance analysis of the queuing system encompasses the evaluation of queue length and waiting time.

2.2 Automobile queue optimization using response surface

The Response Surface Methodology (RSM) has numerous benefits, including high precision, predictive capabilities, the ability to identify the optimal point, and the ability to detect interactions between factors and the relationship between independent variables and the response. The usage of RSM is widespread in both manufacturing and services industries. The factors employed can

encompass both qualitative and quantitative aspects. Additionally, RSM optimizes the queue system performance. The factors used are service time, number of servers, and queue discipline. Responses that are commonly used are queue length, production rate, queue cost, and waiting time [48].

Researchers often combine RSM with several other methods to determine the optimum point. Researchers combine RSM with numerical analysis to analyze Markov chain models of class-based queuing systems (CBQ). First- and second-order models are used for analysis [49]. RSM is used to optimize customer waiting time and queue length. Experiments are carried out using a simulation approach [50]. The optimum poisson rate value that minimizes the number of served passengers and remaining passengers is found using RSM. The Poisson value is used to estimate the characteristics of the queue system [51]. The queue system is approached with several models. RSM is used to select a queue model that has superior performance. The responses analyzed include queue length, customer waiting time, departure rate, and productivity [52]. Optimum conditions are determined based on the values of factors that influence response, such as the number of servers and workers. RSM is used to determine the number of servers and workers, which optimizes sales, queue length, and queue cost. RSM was analyzed using a central composite design (CCD). CCD is an optimization technique that involves observing the optimum area. Several observed points include factorial points, axial points, and central points [53].

RSM is commonly integrated with experimental design. The purpose of experimental design is to determine various factors that impact the responses. The factors that contribute to responses in RSM are used to identify the optimal queue performance point. The evaluation of queue performance is based on the average customer waiting time and service time [54]. There are 29 RSM and DoE application industry case studies discovered, spanning the healthcare, retail, logistics, educational, marketing, after-sales, and catering industries. RSM and DoE applications have not yet been used in queue systems, particularly in automobile workshops [55].

2.3 Research gaps

Based on the literature review, there are several gaps, specifically:
1. The utilization of RSM to enhance the efficiency of automobile workshop queue system has not been put into practice as of now. Recent study has solely discovered precise experimental design models and variables that affect the performance of automobiles queue. [56].
2. Several prior research utilized CCD to ascertain the optimum point. An inherent limitation of CCD is the absence of discrete values for the axial points. Queuing system experiments revolve around discrete elements.
3. Previous studies employed a simulation methodology to conduct experiments. The method of direct observation has not been extensively implemented thus far.
4. No linear equations capable of predicting optimal queue performance have been discovered.
5. The significance of using arrival rate and interarrival time in RSM analysis has not been thoroughly examined. For instance, the time between the arrival of customers may vary even when the rate of arrival remains constant. Reduced interarrivals will result in extended waits and impact the efficiency of the queuing system.

This study encompasses several primary objectives. The first objective of this study is to optimize automobile workshop queue performance using RSM. Optimization is carried out using the box-behnken design, which uses discrete factors. Experiments are conducted by integrating numerous factors and implementing them in an actual queuing system. The second purpose is to establish the correlation between independent variables and responses and derive an equation that can predict the optimal point. The third objective aims to assess the impact of arrival rate and interarrival time on predicting the optimal queue performance. This investigation is being conducted at an SUV repair facility with a daily capacity of 50 vehicles. Routine maintenance services are the subject of this study. The workshop is available seven days a week and eight hours each day. The data collection period spanned a duration of three years to ensure sufficient and reliable results. The queuing discipline employed is based on the principle of first-come, first-served and does not incorporate any form of prioritization.

3. Experimental Procedure

3.1. Factors

Previous study has identified factors that impact the automobile workshop queue system [56]:
1. Number of servers. An analysis is required to determine the optimal value of the number of servers, as it directly impacts queue performance.
2. Number of phases. A worker carries out all

stages of routine maintenance services in one phase. When multiple phases are used, the activity is divided into several stages. If the task is divided into three phases, each phase will include the following activities: The first phase consists of examining the brake pads and tires, the oil leak, the ball joint, the tie rod, the wheel bearing, and the shock absorber leak. The second phase consists of examining radiator water, windshield washer fluid, windshield wiper rubber, brake fluid, exterior and interior lighting, engine and body electricity, air conditioning units, and the battery's performance. The third phase consists of engine tuning.

3. Number of workers per phase. The number of workers affects queue performance. If the number of workers increases, the number of customers served also increases, reducing the number of leaving customers.

4. Arrival rate and interarrival time. The arrival rate depends on the number of customers who visit daily. Interarrival time is the duration of the arrival between customers. These two factors will be analyzed for their accuracy in determining the optimum point.

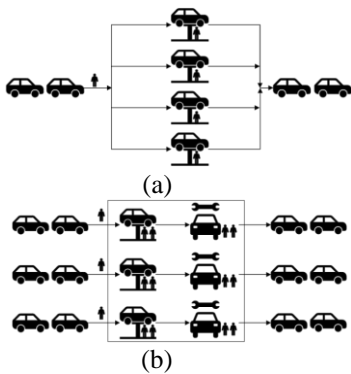5. Layout Type. There are two types of layouts as shown by figure 1.

(a)

(b)

Fig. 1. Layout type

Figure 1a demonstrates that Type A is extensively utilized. Upon customers parking their cars in front of the customer service area, service agents record all repairs and generate job cards. The customer proceeds to the workshop area, and subsequently continues to the waiting room. The Type B design, as shown in Figure 1b, enables customers to remain inside the vehicle. Positioning the service agents adjacent to the customer eliminates the need for the customer to exit the vehicle. The term "layout" refers to a category factor. While the assignment of values to categorical factors may lack a clear logic, it is typical to assign the values of -1 and 1 to type A and type B,

respectively, in order to enhance interpretation [57].

6. Worker experience (z_2). Worker experience is a categorical factor. Employees with one year of work experience are considered to have low experience. High-experienced workers are those who have more than three years of relevant experience.

7. Server area. An adequate server area will support the workers' performance, thereby increasing queue performance.

3.2. Responses
There are several responses analyzed:

1. Queue cost. The queue cost is divided into two categories: capacity costs and customer waiting costs. Capacity costs consist of electricity costs per day, worker costs per day, equipment depreciation per day, and equipment maintenance costs per day. Customer waiting costs include the cost associated with waiting space, which pertains to the extent of the driveway allocated for automobiles awaiting service and the financial impact incurred by the loss of a customer or the opportunity cost resulting from a customer's refusal to wait. Queue cost is calculated in Indonesian Rupiahs (IDR).

2. Service time. Service time refers to the amount of time required to repair one car. The consideration of service time is a crucial response in the development of service designs [58].

3. The average customer waiting time is the average time required by customers to wait for service and get service.

4. The number of customers. Both the number of customers waiting for service and the number of customers being served contribute to the number of customers in the queue system. Observations of the number of customers are carried out every hour, and the results are averaged.

3.3. Steepest descent
Performing the steepest descent determines the location of the optimal point. The factors utilized have a significant impact on responses and are determined through the process of factor screening. The estimation commences with the present state. The equation model employed is a first-order equation model with two replications, enabling an increase in the degree of freedom. If the impact of the arrival rate on the response is negligible, a comprehensive steepest descent analysis will be conducted for all arrival rates collectively. If the response is impacted by the

arrival rate, the analysis is conducted separately for each arrival rate. Curvature is a sign of proximity to the optimum point. Equations 1 and 2 estimate the interaction that causes curvature.

$$\text{Sum of Squares}_{interaction} = \frac{1}{n2^k}(\text{Contrast}_{interaction})^2 \qquad (1)$$

$$F = \frac{\text{Sum of square}_{interaction}}{\sigma^2} \qquad (2)$$

If the F value is below the predetermined threshold, it is permissible to disregard the impact of the interaction or conclude that there is no potential for curvature. The assessment of curvature is also conducted using F test as shown by equation 3:

$$F = \frac{\frac{n_f n_c (\overline{y_f} - \overline{y_c})^2}{n_f + n_c}}{\sigma^2} \qquad (3)$$

Where:

n_f = Number of observations around the center point

n_c = Number of observations at the center point

$\overline{y_f}$ = Average response around the center point

$\overline{y_c}$ = Average response at the center point

If the F value is lesser than the cut off value, the curvature effect is absent, and the observation is continued.

3.4. Response surface methodology

Second-order RSM is carried out to determine the optimum point. The chosen experimental design is Box-Behnken, as it is suitable for discrete factors. If the arrival rate does not affect the response, a second-order analysis will be performed for all arrival rates. When the arrival rate exerts an influence on the response, it is necessary to do a repeated analysis for each individual arrival rate.

Arrival rate and interarrival time are the two variables that are used in RSM. To predict the values of independent factors under ideal conditions, both factors are classified as outer design. The most optimal point value is determined by simultaneously considering the values of the four responses. The equation for determining the optimal value is shown by equation 4.

Where:

T = Target value

y = Response value

U = Upper value

d_i = Desirability value for each response

$$\mathbf{d}_i = \begin{cases} 1; & \text{if } y < T \\ \left(\frac{U-y}{U-T}\right)^r; & \text{if } T \le y \le U \\ 0; & \text{if } y > U \end{cases} \qquad (4)$$

If the value of the response is lower than the target, the desirability value is assigned as one. If the value of the response exceeds the upper threshold, the desirability value is set to zero. In equation 5, the desirability value for each response is substituted to produce the total desirability value. As the D value approaches one, the overall response values are optimized.

$$D = \left(\prod_{i=1}^{4} d_i\right)^{\frac{1}{4}} \qquad (5)$$

Where:

D = Total desirability

$\mathbf{d}_i$ = Desirability value for each response

4. Results
4.1. Optimization with arrival rate as outer design
4.1.1 Steepest descent

Based on findings in previous research, queue cost and service time are significantly affected by the number of servers, the number of phases, and the number of workers per phase. Responses are also significantly influenced by the arrival rate. The appropriate number of servers, phases, and workers must be determined for each arrival rate. The arrival rate ranges from 21 to 50 cars per day, thereby necessitating the application of the steepest descent analysis within this interval. Previous research findings indicate that the server area does not significantly influence responses. Consequently, the default setting for server space allocation remains at nine square meters. The absence of an impact of worker experience on response necessitates the utilization of workers with little experience for conducting steepest descent analysis. Layout type B is selected due to its detrimental impact on the average customer waiting time, service time, and queue cost [56]. The experimental design consists of three factors, two replications, and four center points. These center points determine the presence of an optimum site. The workshop normally uses two servers, two phases, and two workers for each step. Consequently, the number of servers, phases, and workers ranges from one to three. The first-order analysis is performed when the arrival rate is 21 cars per day. A first-order approach is used to analyse the experimental results. Table 1 shows the findings of the first-order analysis.
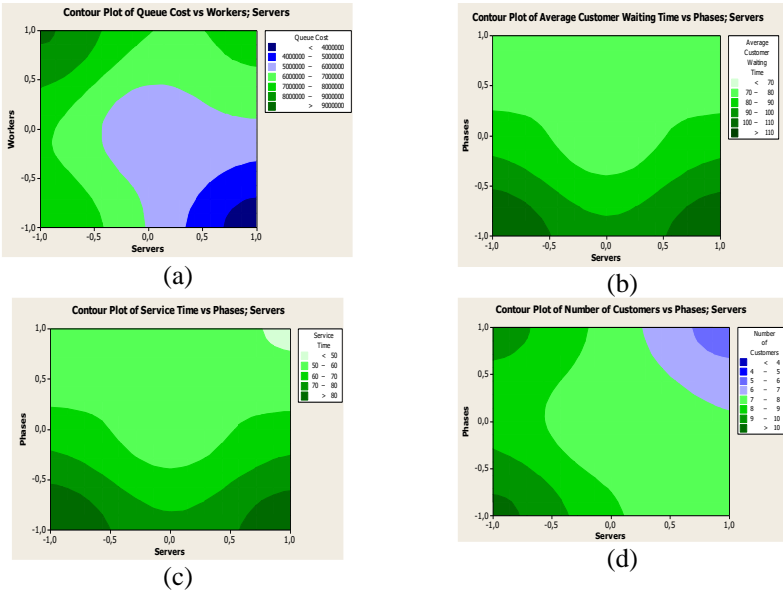
(a)

(b)

(c)

(d)

Fig.2. Contour plot of responses

Tab.1. First order analysis

Term	Queue Cost		Average Customer Waiting Time		Service Time		Number of Customers	
	Coef.	F	Coef.	F	Coef.	F	Coef.	F
Constant	5349333		74.73		55.23		7.38	15.44
Linear		14.89		321.460		498.05		19.94
Servers	-1512833	27.34	-0.55	0.880	-0.88	3.48	-1.60	49.11
Phases	106250	0.13	-18.08	963.290	-18.25	1490.58	-0.63	7.67
Workers	1200000	17.20	-0.26	0.190	-0.13	0.08	-0.40	3.04
Square		6.12		118.210		143.27		1.94
R Square		77.2%		98.6%		99.09%		80.46%

Table 1 reveals that the number of servers, phases, and workers have a linear and quadratic effect on queue cost. This result is indicative of the existence of curvature, with the starting observation point being near the minimum queue cost position. This finding is supported by figure 2a. Figure 2a shows that the queue cost value is close to the minimum location. The average customer waiting time is impacted quadratically by the number of phases. This also indicates the proximity to the minimum location as shown by figure 2b. The number of phases has a quadratic effect on service time. The observation point is also near the minimum service time as shown by figure 2c. The number of customers is influenced linearly by the number of servers, the number of phases, and the number of workers. Observations at the origin are near the minimum point, as shown in Figure 2d. According to first-order analysis, only the number of phases influences the service time. However, based on the previous study [56], Both the number of workers and the number of phases have a negative effect on service time and the average waiting time for customers. Thus, the steepest descent approach involves the utilization of factors and begins with a large initial value. The findings derived from the steepest descent analysis are illustrated in Figure 3.
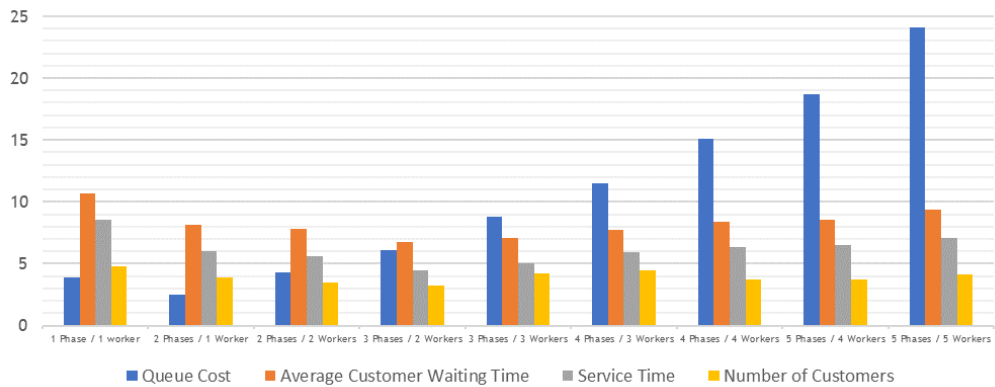
Fig. 3. Steepest descent analysis for service time and average customer waiting time.

The steepest descent analysis for service time and average customer waiting time, as depicted in Figure 3, can identify not only the optimal service time but also the minimum values for average customer waiting time and number of customers. The minimum value for both responses occurs when there are three phases and two workers. The increase in both factors leads to an increase in the response value. At three phases and two workers, the number of customers reaches its minimum before exhibiting fluctuating patterns. This phenomenon occurs due to the escalating service durations, resulting in the departure of certain customers from the queue.
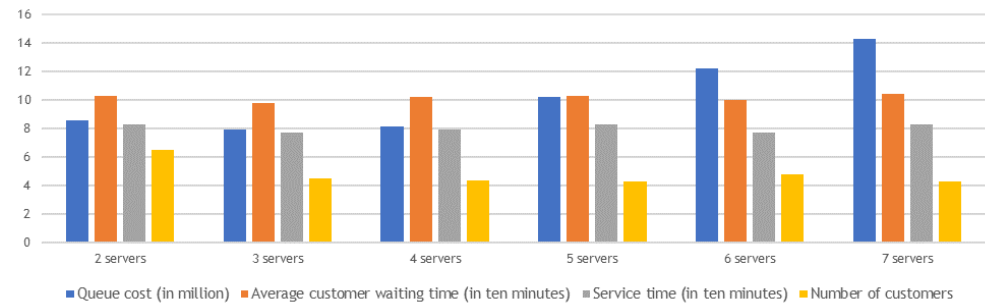

Fig. 4. Steepest descent analysis for queue cost

The subsequent step is to determine the optimal queue cost, as shown in Figure 4. First-order analysis demonstrates that the number of servers has a negative impact on the queue cost. The steepest descent begins from the large value to the small value. The decrease in the number of servers and phases is not possible because it would result in an increase in service time. The steepest descent method is executed with three phases and two workers. The minimum value for queue cost is attained when there are three servers. Increasing the number of servers results in a reduction in queue cost to a minimum threshold, after which the cost begins to escalate because capacity costs are greater than customer waiting costs. As the number of servers increases, the number of customers decreases proportionally. The number of servers does not have any effect on the service time or the average customer waiting time. The next response surface procedure is conducted on a set of three servers, involving three steps and employing two workers.

4.1.2 Optimization
The utilization of the second-order model is employed for the purpose of optimization. The chosen design is the box-behnken design, as indicated in Table 2, due to the discrete characteristics of every factor involved. Each group is replicated four times.

Tab.2. Box–behnken design for three factors

	Treatment		
	1	2	3
Group 1	x	x	
Group 2	x		x
Group 3		x	x

There are three center points. The total number of experiments is 15. The findings of the box-behnken experiments are presented in table 3, specifically focusing on the scenario when the arrival rate is 21 cars per day. According to Table 3, it can be observed that all factors exhibit a quadratic impact on the responses. The observed point is the minimum point. Subsequently, the process of optimization is conducted by employing equations 4 and 5. The optimization outcomes are depicted in Figure 5.

Tab. 3. Second order analysis with box-behnken design

| Terms	Queue cost		Average customer waiting time		Service Time		Number of customers	
	Coef.	F	Coef.	F	Coef.	F	Coef.	F
Linear		174.13		131.13		147.01		22.97
Servers	-269083	5.92	0.281	0.12	10.470	1.87	-123.438	47.33
Phases	1912500	298.97	0.073	0.01	0.3289	0.18	-0.09375	0.27
Workers	1631250	217.50	-15.779	393.25	-160.496	438.97	-0.82813	21.30
Square		18.15		67.63		78.67		3.71
Interaction		21.40		0.71	155.418	1.27	0.86979	2.76
R Square		99.23%		99.17%		99.27%		94.64%

(b)

(c)

(d)

(e)

(a)
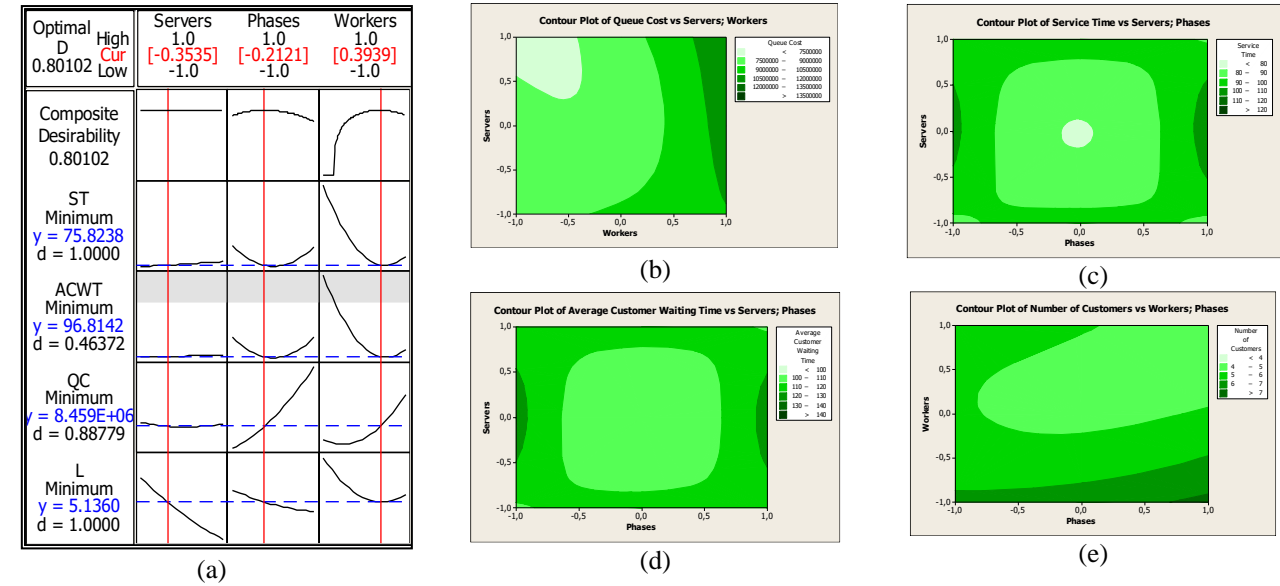
Fig. 5. Optimization

Figure 5a depicts that at the optimal or minimum point, the values for the number of servers, number of phases, and number of workers are three, three, and two, respectively. The minimum values for queue cost, service time, average customer waiting time, and number of customers are determined to be IDR 8,458,723, 76 minutes, 97 minutes, and 5 cars, respectively. According to Figure 5b, when the number of workers is adjusted to one, the queue costs are minimized, whilst setting it to two maximizes the desirability function. Figures 5c, 5d, and 5e show that the minimum values for

service rate, average customer waiting time, and number of customers are in the correct location and maximize the desirability function. The composite desirability value of 0.80 indicates that the estimate closely approximates the optimal position. The optimizations are conducted uniformly for an arrival rate ranging from 22 to 50 automobiles per day, and the results are displayed in Table 4. QC, ST, ACWT, and N refer to queue cost (in million), service time (in minutes), average customer waiting time (in minutes), and number of customers, respectively.

Tab.4. Optimization for arrival rate between 22 and 50 cars

No	Arrival	Factor			Optimum Condition					Ratio
		Servers	Phases	Workers	QC	ST	ACWT	N	Desirability	
1	21	2.65	2.79	2.39	8.46	76	97	5	0.80	1.26
2	22	2.00	2.83	2.29	8.74	77	98	5	0.79	1.76
3	23	3.00	3.00	2.00	9.98	75	96	6	0.79	1.01
4*	24	5.00	4.00	2.00	10.49	78	100	9	0.78	1.08
5	25	3.79	2.85	2.29	9.39	74	97	5	0.79	1.02
6	26	4.00	2.67	2.41	9.94	75	99	5	0.76	1.02
7	27	4.34	2.75	2.29	10.63	76	100	6	0.73	0.98
8	28	4.48	3.47	2.45	15.03	76	101	5	0.71	0.99
9	29	6.00	2.75	2.33	11.81	73	98	5	0.76	0.74
10*	30	3.31	2.97	1.97	5.93	46	70	6	1.00	0.87
11	31	4.61	2.79	2.39	12.07	75	101	6	0.74	1.05
12	32	4.93	2.79	2.37	12.24	76	101	6	0.74	1.03
13	33	5.01	2.85	2.31	12.31	74	100	6	0.74	1.02
14*	34	8.00	2.59	2.01	14.16	104	130	6	0.89	0.92
15	35	6.62	2.75	2.33	13.15	75	103	6	0.71	0.83
16	36	6.23	2.77	2.33	13.04	74	102	6	0.71	0.89
17*	37	7.23	3.23	1.53	11.13	90	117	7	0.82	0.96
18*	38	3.92	2.91	1.53	4.39	38	68	6	1.00	0.77
19	39	6.18	3.00	2.15	15.00	76	106	6	1.00	1.00
20	40	6.83	2.75	2.39	14.99	76	106	7	0.73	0.93
21*	41	4.00	2.65	2.02	5.35	37	69	6	1.00	0.79
22	42	7.15	2.79	2.35	15.80	77	108	7	0.69	0.94
23	43	7.50	2.87	2.35	16.52	77	108	7	0.59	0.92
24	44	7.88	2.75	2.37	16.82	77	109	7	0.59	0.90
25	45	7.69	2.79	2.33	17.07	77	109	7	0.65	0.94
26*	46	-	-	-	-	-	-	-	-	-
27	47	9.00	2.85	2.33	18.40	78	111	8	0.50	0.85
28*	48	6.88	2.85	2.25	12.80	55	89	8	0.76	0.80
29*	49	5.85	2.73	2.31	13.13	56	91	8	0.81	0.98
30	50	8.39	2.87	2.31	18.57	76	112	7	0.61	0.94

Table 4 reveals that at three phases, two workers, and a certain number of servers, the queue cost reaches its lowest point. The number of servers increases proportionally with the arrival rate. There are aberrations; however, using three phases and two employees per phase does not yield the same optimal service time. For instance, the ninth experiment is conducted under the condition of an arrival rate of 29 cars per day. The optimal service time per car is 73 minutes. There are a total of six servers. The tenth experiment is conducted when the daily arrival rate was 30 cars. The optimal server time per car is 46 minutes, and three servers are required. If the arrival rate increases, the number of servers should also increase. At an arrival rate of 30, however, the number of servers decreases. This occurs because service time decreases, allowing each server to serve more cars, and the number of servers decreases.

The thirteenth experiment is conducted when the daily arrival rate is 33 cars. The optimum service time per car is 74 minutes. The required number of servers is five. The fourteenth experiment is conducted with an arrival rate of 34 cars per day. The service time is 104 minutes per car. A total of eight servers is expected. The fourteenth experiment requires a greater number of servers compared to the thirteenth experiment. This occurs as the service time increases, reducing the number of cars served while increasing the number of servers to anticipate the arrival rate. Experiments 18, 21, 28, and 29 demonstrate the same phenomenon. Consequently, the arrival rate has no significant effect on the optimum queue cost. The rightmost column represents the service utility, which is the ratio of the arrival rate to the number of cars successfully served. The service utility value approaches one under optimal conditions. This implies that the arrival rate and the service capacity must be equal at the optimal point. Table 4 presents findings that are in contradiction to prior studies that examined queuing systems through a probabilistic framework. The probability technique utilizes the

average arrival rate and service time to determine the optimal number of servers. The number of servers remains constant across all arrival rates. The results of this research are contradictory. The number of servers fluctuates by arrival rate and service time. Adjusting the number of phases and the number of workers can minimize the service time and average customer waiting time, although the value may differ. Therefore, there is a disparity in the number of servers, despite the arrival rate being nearly identical.

Table 4 shows that the smallest queue cost is achieved when the service time is at its lowest. Experiment 20 is conducted under the condition of an arrival rate of 40 automobiles per day. The minimum service time for each automobile is 76 minutes. The minimum queue cost is IDR 14,987,257. Experiment 21 is conducted under the condition that the arrival rate is observed to be 41 automobiles per day. The optimum service time is determined to be 37 minutes. The minimum cost of the queue is IDR 5,350,486. The queue cost in experiment 21 are expected to be similar to those observed in experiment 20. Queue costs are lower in experiment 21 because the optimum service time is very low, allowing more customers to be served while lowering queue cost. The optimum average customer waiting time is achieved at 3 phases and 2 workers. The optimum number of customers is obtained when the service utility

equals one, and its magnitude is proportional to the number of servers and the arrival rate. One noteworthy aspect pertains to Experiment 18. Experiment 18 is conducted when the daily arrival rate is 38 cars. The optimal service time is the minimum value compared to the other service times. Consequently, the optimal queue cost is observed to be the lowest compared to the other experiments.

The third experiment is carried out when the arrival rate was 23 cars per day. The optimum values for queue cost, service time, average customer waiting time, and number of customers are IDR 9,983,276; 75 minutes per car; 96 minutes per car; and 6 cars. The fourth experiment is carried out when the arrival rate is 24 automobiles per day. The optimum values for queue cost, service time, average customer waiting time, and number of customers are IDR 10,485,954, 78 minutes per automobile, 100 minutes per automobile, and 9 automobiles. The optimum value of the fourth experiment should be close to the optimum value of the third experiment. The third experiment has nearly the same optimum service time as the fourth experiment, but the fourth experiment has a higher queue cost and a larger number of customers. The reason for this is because the interarrival time between customers in the fourth experiment is shorter than in the third. This is illustrated in Figure 6.
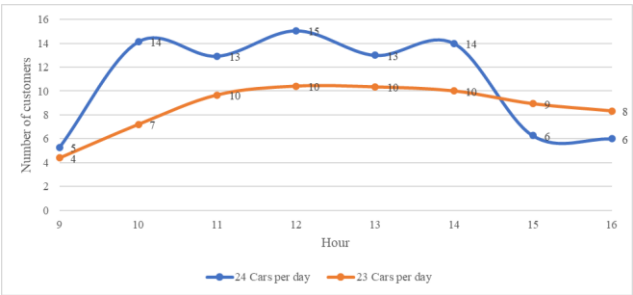

Fig. 6. Number of customers

Figure 6 depicts the number of customers in each hour. Observations are carried out from 9 a.m. to 4 p.m. Figure 6 shows that the arrival time between customers at an arrival rate of 24 cars per day is greater. Initially, there are 5 customers at 9 a.m. and 14 customers at 10 a.m., therefore there are 9 customers arriving between 9 and 10 a.m. This leads to many customers in the queue and higher queuing costs. The arrival rate is ineffective at improving queue performance, hence the

interarrival time must be addressed.

Figure 7 depicts the relationship between queue cost and number of servers at arrival rates of 45 (experiment 25) and 46 (experiment 26). Experiments with an arrival rate of 45 autos each day yield a profit of IDR 450,000 per customer. The experiments on an arrival rate of 46 cars per day are carried out by providing discounts to customers so that customers are only charged a service cost of IDR 250,000.
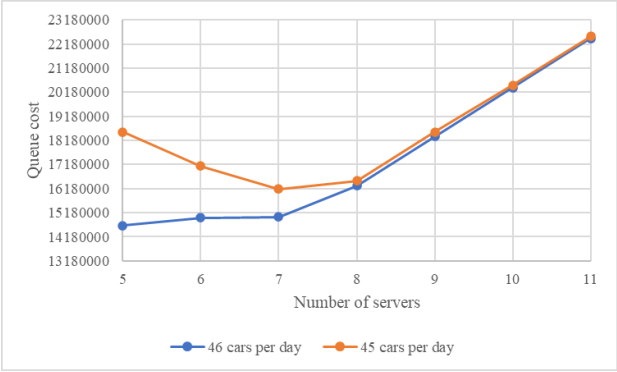
Fig. 7. Queue cost

The queue cost of experiment 25 can be optimised. The queue costs in experiment 26 cannot be optimized, hence the optimum values for service time, average customer waiting time, and number of customers cannot be established concurrently. The optimum number of servers in experiment 26 is less than 5 servers, which is not attainable since the number of customers who do not receive service and the number of customers who abandon the waiting system will increase. Differences in cost composition cause the optimal value of the overall response to fluctuate. Determining the optimal value based on queue cost is inaccurate.
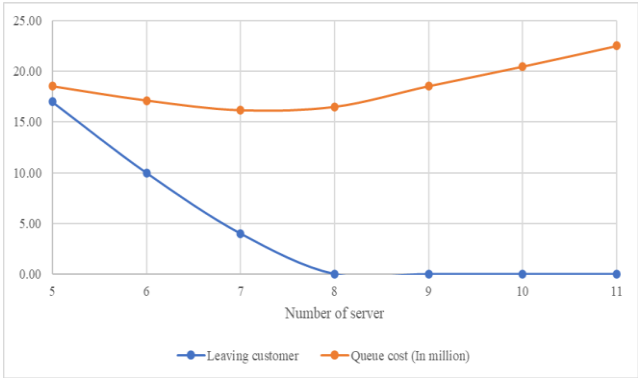

Fig. 8. Comparison between leaving customers and queue cost

Figure 8 shows a graph of the number of leaving customers and the queue cost versus the number of customers. Customers who enter the queue system and subsequently depart are referred to as leaving customers. The queue cost is minimized when there are no customers departing. The number of leaving customers is a more precise indicator for calculating the optimal value. Based on this, experiments are carried out by leaving customers as a response. The experiment started with screening analysis, steepest descent, and optimization. The factors analysed are the number of servers, server area, number of phases, number of workers, interarrival rate, layout type, and worker experience. The responses analysed are number of leaving customers, service rate, average customer waiting time, and number of customers.

4.2. Optimization with interarrival time as outer design

4.2.1 Screening analysis

The full design for screening analysis includes the application of the interarrival time as a factor. The design used is fractional factorial 2_{IV}^{7-2} and is presented in Figure 9. This design is a four-resolution design with the defining relation I = ABCDF = ABDEG = CEFG. A single factor correlates with a four-factor interaction. Two factors of interaction correlate with three factors of interaction. Although there are correlations between interactions, the design remains orthogonal. The design comprised 32 experiments, each of which is reproduced twice, resulting in a total of 64 experiments. Replicating the 32 experiments results in an increased degree of freedom, which enables the estimation of all coefficients. The screening factor results are displayed in Table 5.
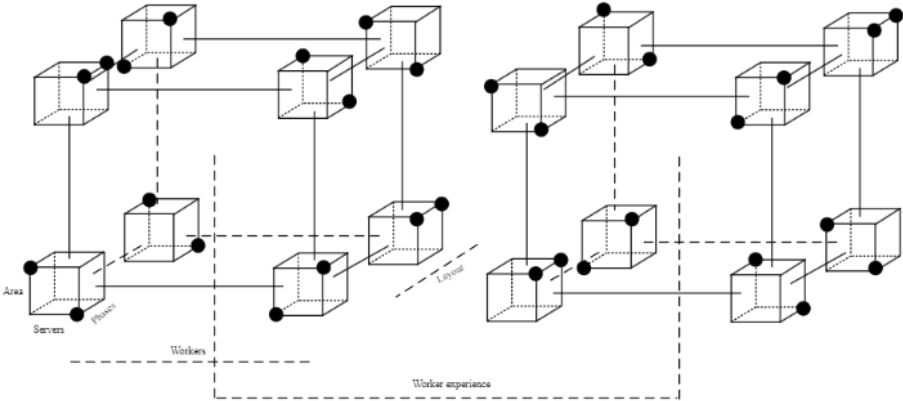
Fig. 9. Fractional factorial 2_{IV}^{7-2}

Tab. 5. Optimization for interarrival time between 9 to 20 minutes

Term	Leaving customer		Service time		Average Customer Waiting Time		Number	
	Effect	P	Effect	P	Effect	P	Effect	P
Servers	-25.65	0	0.625	0.046	0.1	0.772	-0.971	0
Area	-0.02	0.884	0.131	0.667	-0.09	0.794	0.037	0.109
Phases	-1.65	0	-4,123	0	-4.57	0	-0.079	0.001
Workers	0.15	0.349	-3,549	0	-3.62	0	0.007	0.748
Interarrival Time	-18.71	0	0.284	0.353	-4.49	0	-3,399	0
Layout	-0.1	0.518	0.045	0.882	-25.24	0	-0.002	0.922
Experience	0.09	0.587	-0.149	0.625	-0.47	0.189	-0.019	0.409
Servers*Interarrival	4.16	0	-0.611	0.05	-0.2	0.576	0.186	0
R square	99.91%		90.63%		99.34%		99.85%	

Table 5 presents the results of the screening analysis. The number of servers, phases, and interarrival time have a negative effect on the number of leaving customers. Increasing the number of servers and the number of phases reduces the number of customers leaving the queue system. A decrease in the interval between the customer arrivals, or a decrease in the interarrival time, results in an increase in the number of customers departing. Smaller interarrivals cause customers to accumulate in lines, increasing the likelihood that customers may leave the queuing system. Server-interarrival time interaction has a positive effect on the number of leaving customers. This interaction demonstrates a curve in the number of leaving customers. Both the number of phases and the number of workers have a detrimental impact on service time. Increasing the number of phases and workers will reduce service time.

The average customer waiting time is negatively affected by the number of phases, number of workers, interarrival time, and layout. Increasing the number of phases and workers shortens service time and will minimize customer waiting time. When there is a long interval between customers arriving or an extended interarrival time, the number of customers waiting in line will be reduced, and the likelihood of customers abandoning the queue will be decreased. The layout has a negative effect, so that type B layout shortens the duration of customers queuing.

The number of servers, the number of phases, and the interarrival time have a negative effect on the number of customers. Server-interarrival interaction has a positive effect on the number of customers. Increasing the number of workers and phases reduces service time and average customer waiting time, resulting in a smaller number of customers queuing. An increase in servers reduces the number of customers in queue and gives customers more opportunities to obtain service. Worker experience and server area have no impact on the four responses, and this result is in line with the previous analysis.

4.2.2 Steepest descent

Experiments are carried out at each interarrival time or the interarrival time is set as the outer design. Optimizing interarrival time is impossible due to its unpredictability, but in this experiment, interarrival time is controlled through a booking service mechanism. The steepest descent experiment is carried out using a type B layout and

highly experienced workers.

Steepest descent is carried out to predict the number of leaving customers. The number of phases and workers is specified as three phases

and two workers. The steepest descent analysis is conducted for each vehicle at a nine-minute interarrival time, and the outcome is illustrated in Figure 10.
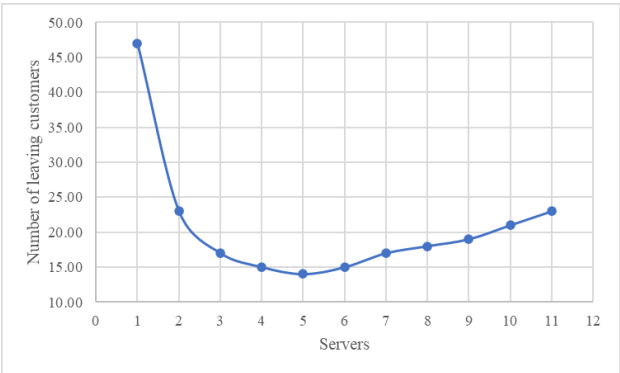

Fig. 10. Steepest descent for losing customers

Figure 10 shows that the relationship between the number of leaving customers and the number of servers is reciprocal. Based on the results of the screening analysis, the equation to predict the number of leaving customers is estimated as follows:

$$L = 0.29x_1x_4 + \frac{51.82}{x_1} + \frac{47.31}{x_4} - 2.36x_4 - 1.57x_2 + 12.58 \qquad (6)$$

Where:

x_1 = Number of servers
x_2 = Number of phases
x_4 = Interarrival time
L = Number of leaving customers

Number of servers and interarrival time are inversely proportional to the number of leaving customers. As the interarrival time and number of servers increase, the number of leaving customers decreases, implying that these two factors have a negative impact on the number of leaving customers. The interaction effect demonstrates the presence of a curve or optimum value for the number of leaving customers. Partial differentiation of equation 6 with respect to the number of servers yields equation 7.

$$\frac{\partial L}{\partial x_1} = 0.29x_4 - \frac{51.82}{x_1^2} \qquad (7)$$

The optimum number of leaving customers is determined by Equation 8 when $\frac{\partial L}{\partial x_1} = 0$.

$$x_1 = \sqrt{\frac{51.82}{0.29x_4}} \qquad (8)$$

The optimum value of the number of leaving customers changes and depends on the value of the interarrival time. The value of the optimal number of servers will decrease as the interarrival time increases. This equation will be compared with the optimization results using the response surface. Based on the Figure 10, the number of leaving customers reaches the minimum value when the number of servers is 5. This value is used for response surface analysis.

Figure 11 shows the findings of the steepest descent analysis for service time and average customer wait time. The correlation between the number of phases and workers is quadratic, leading to an optimal value. The optimal values are achieved by utilizing three servers and two workers.

Fig. 11. Steepest descent for service time and average customer waiting time

The steepest descent experiment yields equation 9 for predicting service time:

$$ST = 0.43x_2^2 - 2.59x_2 + 0.355x_3^2 - 1.36x_3 + 7.72 \qquad (9)$$

Where:
x_2 = Number of phases
x_3 = Number of workers
ST = Service time
The optimal service time value is unaffected by the interarrival time and solely determined by the number of workers and servers. Equations 10 and 11 result from the partial differentiation of equation 9.

$$\frac{\partial ST}{\partial x_2} = 0.86x_2 - 2.59 \qquad (10)$$

$$\frac{\partial ST}{\partial x_3} = 0.71x_3 - 1.36 \qquad (11)$$

Equations 10 and 11 demonstrate that at 3.01 phases and 1.92 workers, the service time value is at its lowest. But the lowest amount varies based on the type of vehicle, the type of service, and other factors.

Equation 12 that predicts the average customer waiting time is derived from the experiment with the steepest descent.:

$$ACWT = 0.42x_2^2 - 2.53x_2 + 0.38x_3^2 - 1.47x_3 - 0.02x_4 + 8.907 \qquad (12)$$

Where:
x_2 = Number of phases
x_3 = Number of workers
x_4 = Interarrival time
ACWT = Average customer waiting time
The service time equation and the average customer waiting time equation are nearly identical. The number of phases, workers, and interarrival time all affect the optimal value of the average customer waiting time. Equations 13 and 14 are produced if the equation is differentiated.

$$\frac{\partial ACWT}{\partial x_2} = 0.84x_2 - 2.53 \qquad (13)$$

$$\frac{\partial ACWT}{\partial x_3} = 0.76x_3 - 1.47 \qquad (14)$$

The optimum value of average customer waiting time was achieved at 3.01 phases and 1.93 workers. This optimum condition is the same as the optimum service time condition. However, the value of average customer waiting time is negatively influenced by interarrival time.
The following is an approximated equation for predicting the number of consumers based on the experiment with the steepest descent.

$$N = \frac{2.24}{x_1} + \frac{0.55}{x_2} + \frac{75.4}{x_4} + 0.02x_1x_4 - 3.04 \qquad (15)$$

Where:
x_1 = Number of servers
x_2 = Number of phases
x_4 = Interarrival time
N = Number of customers
Number of servers, number of phases and interarrival time are inversely proportional to number of customers. The interaction effect also indicates a curvature in the number of customers. However, optimization of the interarrial rate with a value of 9 minutes is carried out on 5 servers, 3 phases, and 2 workers.

4.2.3 Optimization
The optimization of the 9-minute results yields an optimal configuration consisting of 5 servers, 4 phases, and 1 worker. The optimal values for each response are as follows: 12.76 for the number of leaving customers, 34.37 minutes for service time, 43.59 minutes for average customer waiting time, and 6.33 for the number of customers in the queue system.
Additionally, optimization is conducted for various interarrival times using the same method, and the outcomes are displayed in table 6.

Tab. 6. Optimization for interarrival time between 9 to 20 minutes

No	Inter-arrival	Factors (by response surface)			Factors (by equations)			Optimum Condition				
		Servers	Phases	Workers	Servers	Phases	Workers	L	ST	ACWT	N	Desirability
1	9	5.21	4.00	1.00	4.46	3.01	1.93	12.76	34.37	43.59	6.33	0.84
2	10	5.39	3.52	1.63	4.23	3.01	1.93	18.82	26.83	37.81	6.96	0.94
3	11	4.55	4.00	1.38	4.03	3.01	1.93	11.31	26.83	37.81	6.96	0.93
4	12	4.31	3.49	1.87	3.86	3.01	1.93	9.87	26.22	35.03	4.80	0.89
5	13	3.95	3.47	1.00	3.71	3.01	1.93	10.58	25.96	34.29	4.51	0.88
6	14	4.37	3.19	1.91	3.57	3.01	1.93	9.23	25.31	33.01	4.11	0.92
7	15	4.00	3.29	1.97	3.45	3.01	1.93	6.04	25.06	32.34	3.76	0.99
8	16	3.92	3.11	1.87	3.34	3.01	1.93	5.99	24.57	31.97	3.56	0.99
9	17	4.00	3.00	2.00	3.24	3.01	1.93	3.77	25.00	31.05	3.28	1.00
10	18	4.00	2.93	2.05	3.15	3.01	1.93	1.47	24.12	31.81	3.12	0.99
11	19	2.91	2.99	2.01	3.07	3.01	1.93	1.37	25.08	30.49	2.89	1.00
12	20	2.12	3.19	1.87	2.99	3.01	1.93	0.00	24.72	29.96	2.67	1.00

According to Table 6, an increase in the interarrival time value leads to a decrease in the number of leaving customers.

A higher interarrival time leads to a lower number of customers in the queue, resulting in a lower probability for customers leaving the queue system. This is demonstrated by the lower average number of customers waiting in line. The number of phases and workers is constant at each interarrival rate. This is in accordance with equations 10, 11, 13, and 14, which assert that the optimal value of service time and the average customer waiting time are solely determined by the number of phases and workers.

Table 6 demonstrates that equations 8, 10, 11, 13, 14, and 15 yield estimations for the number of servers, phases, and workers that are closely consistent with the outcomes of optimization using the response surface. This equation is applicable to various other interarrival times. Nevertheless, this equation is exclusively utilized for predicting queue conditions for routine maintenance service sections. The mathematical equations that predict queuing conditions for different types of service must be recalculated.

5. Conclusion

The findings of the investigation indicate that several factors, such as the number of servers, phases, workers, arrival rate, and layout significantly influence the responses. The process of optimization is conducted for every arrival rate to determine the optimum number of responses associated with each arrival rate. The optimum conditions are attained when the service utility is one or when the service capacity is equal to the arrival rate, resulting in minimum values for the responses. The optimal value for all responses is mostly influenced by the number of phases and the

number of workers. The service time is minimized by the number of phases and workers, which contributes to the reduction of queue cost.

Nevertheless, there are several deviations. The queue cost, initially used as a response, cannot serve as a reference for identifying optimal conditions because distinct cost structures will yield various optimal values at the same arrival rate. The analysis results indicate that the optimal queue cost value is equivalent to the optimal number of leaving customers. This occurs because the optimal cost structure is achieved when the number of leaving consumers is minimized. The cost of providing servers is equivalent to the cost of queueing. These findings provide a significant contribution to the optimization of queuing systems. The arrival rate is not relevant for analysis as the inter-arrival rate amongst customers can vary despite having the same arrival rate. This study proposes that the interarrival time is a more relevant factor to consider along with analyzing and optimizing queue performance.

This study presents a comprehensive analysis of the relationship between factors and responses. The number of servers and interarrival time are inversely proportional to the number of leaving customers and number of customers so this is in accordance with several previous studies which stated that the interarrival rate and arrival rate have an exponential distribution. The optimal values of service time and average customer waiting time are dependent upon the number of phases and workers. However, the value of average customer waiting time is directly proportional to interarrival time. An increase in the interarrival time will lead to a decrease in the average customer waiting time. This study establishes a foundation for researchers to further minimize service time and enhance service stability. In addition, it is important to

conduct additional study that examines mathematical equations for various kinds of services.

This study offers managerial insights for automobile workshop managers to prioritize factors that influence service time, as service time directly affects queue performance. In addition, it is crucial to carefully monitor and effectively manage the interarrival time to minimize customer attrition.

References

[1] S. A. Mangkona and I. Murdifin, 'Implementation of Queue Model for Measuring the Effectiveness of Suzuki Car Maintenance', *World Journal of Business and Management*, Vol. 3, No. 1, (2017), p. 55.

[2] H. Q. Nguyen and T. Phung-Duc, 'Strategic customer behavior and optimal policies in a passenger–taxi double-ended queueing system with multiple access points and nonzero matching times', *Queueing Syst*, Vol. 102, No. 3, (2022), pp. 481-508.

[3] S. Srivastava, 'Queuing theory in workshop', *International Journal of Science, Technology & Management*, vol. 04, no. 01, (2015), pp. 88-95.

[4] D. Bannikov, N. Sirina, and A. Smolyaninov, 'Model of The Maintenance And Repair System In Service', *Transport Problems*, Vol. 13, No. 3, (2018), pp. 1-10.

[5] S. Vijay Prasad, R. Donthi, and M. K. Challa, 'The sensitivity analysis of service and waiting costs of a multi server queuing model', in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, (2020).

[6] K. State, A. Adeniran, and M. S. Burodo, 'Application of Queuing Theory and Management of Waiting Time Using Multiple Server Model : Empirical Evidence From Ahmadu Bello', *International Journal of Scientific and Management Research*, Vol. 5, No. 4, (2022), pp. 159-174.

[7] D. Grozev, M. Milchev, and I. Georgiev, 'Study the work of specialized car service as queue theory', *International Scientific Journal 'Mathematical Modeling'*, Vol. 34, No. 1, (2020), pp. 31-34.

[8] A. Michael K., A. Saheed A., and A. Awaw K., 'Congestion Problem during Covid-19 in the University College Hospital , Ibadan , Oyo State , Nigeria : An Application of Queuing Theory', *International Journal of Mathematics and Statistics Studies*, Vol. 11, No. June, (2023), pp. 61-66.

[9] S. Aalto and Z. Scully, 'Minimizing the mean slowdown in the M/G/1 queue', *Queueing Syst*, Vol. 104, No. 3, (2023), pp. 187-210.

[10] S. D. Soorya and K. S. Sreelatha, 'Application of Queuing Theory to Reduce Waiting Period at ATM Using A Simulated Approach', in *IOP Conf. Series: Materials Science and Engineering*, (2021), pp. 1-8.

[11] B. H. Margolius, 'The periodic steady-state solution for queues with Erlang arrivals and service and time-varying periodic transition rates', *Queueing Syst*, Vol. 103, No. 1, (2023), pp. 45-94.

[12] S. Robinson, 'Modelling without queues: adapting discrete-event simulation for service operations', *Journal of Simulation*, Vol. 9, No. 3, (2015), pp. 195-205.

[13] J. Dehantoro, D. Sumiardi, and O. Hijuzaman, 'Analysis of Vehicle Service Queuing System Using Arena in Authorized Workshop', *International Journal of Science and Research (IJSR)*, Vol. 5, No. 5, (2016), pp. 112-117.

[14] M. Kumar Rajuwar and D. Kalita, 'Simulation of queuing System for Car Service Center using Arena Simulation Software', *International journal of Production Engineering*, Vol. 4, No. 2, (2018), pp. 1-12.

[15] Y. Barlas and O. Özgün, 'Queuing systems in a feedback environment: Continuous versus discrete-event

simulation', *Journal of Simulation*, Vol. 12, No. 2, (2018), pp. 144-161.

[16] E. Harahap, D. Darmawan, Y. Fajar, R. Ceha, and A. Rachmiatie, 'Modeling and simulation of queue waiting time at traffic light intersection', *J Phys Conf Ser*, Vol. 1188, No. 1, (2019).

[17] Vipin Kumar Solanki, S. Srivastava, and S. Shrivastava, 'Case Study for Bank ATM Queuing Model', *Turkish Online Journal of Qualitative Inquiry (TOJQI)*, Vol. 7, No. 12, (2021), pp. 6878-6884.

[18] N. Zychlinski, 'Applications of fluid models in service operations management', *Queueing Syst*, Vol. 103, No. 1, (2023), pp. 161-185.

[19] T. M. Abuhay, S. Robinson, A. Mamuye, and S. V. Kovalchuk, 'Machine learning integrated patient flow simulation: why and how?', *Journal of Simulation*, Vol. 17, No. 5, (2023), pp. 580-593.

[20] F. Chiacchio, L. Oliveri, S. M. Khodayee, and D. D'Urso, 'Performance Analysis of a Repairable Production Line Using a Hybrid Dependability Queueing Model Based on Monte Carlo Simulation', *Applied Sciences (Switzerland)*, Vol. 13, No. 1, (2023).

[21] U. C. Okonkwo, I. P. Okokpujie, B. N. Odo, and O. S. I. Fayomi, 'Workshop Queue System Modification Through Multi Priority Strategy', *J Phys Conf Ser*, Vol. 1378, No. 2, (2019).

[22] L. Ravner and Y. Sakuma, 'Strategic arrivals to a queue with service rate uncertainty', *Queueing Syst*, Vol. 97, No. 3, pp. 303–341, 2021,

[23] M. W. Isken, O. T. Aydas, and Y. F. Roumani, 'Queueing inspired feature engineering to improve and simplify patient flow simulation metamodels', *Journal of Simulation*, Vol. 00, No. 00, (2023), pp. 1-18.

[24] M. Irisbekova, 'Improving methodology of automobile operating companies activities simulation modeling', *E3S Web of Conferences*, Vol. 05024, No. 05024, (2021), pp. 1-8.

[25] P. Raghuwanshi and A. Goyal, 'Evaluation and Improvement of Plant Inventory and Layout Design in Automobile Service Industries', *Int J Innov Res Sci Eng Technol*, Vol. 4, No. 9, (2015), pp. 8178-8185.

[26] J. E. Sinebe, U. C. E. Okonkwo, and L. C. Enyi, 'Simplex Optimization of Production Mix : A Case of Custard Producing Industries in Nigeria', *Int J Appl Sci Technol*, Vol. 4, No. 4, (2014), pp. 180-189.

[27] H. Almomani and N. Almutairi, 'Vehicles Maintenance Workshops Layout and its Management to Reduce Noise Pollution and Improve Maintenance Quality', *Journal of Environmental Treatment Techniques*, Vol. 8, No. 4, (2020), pp. 1352-1356.

[28] S. Ďutková, K. Achimský, and D. Hoštáková, 'Simulation of queuing system of post office', *Transportation Research Procedia*, Vol. 40, (2019), pp. 1037-1044.

[29] N. M. Ahmed, M. A. El Sharief, and A. B. A. Nasr, 'Implement Lean Thinking in Automotive Service Centers to Improve Customers' Satisfaction', *Int J Sci Eng Res*, Vol. 6, No. 6, (2015), pp. 576-583.

[30] V. P. Kommula, H. Mapfaira, J. Gandure, K. Mashaba, R. Monageng, and O. Samuel, 'Improving Productivity of a Machine Workshop through Facilities Planning', *International Journal of Mining, Metallurgy & Mechanical Engineering (IJMMME) Volume*, Vol. 3, No. 3, (2015), pp. 187-192.

[31] I. Y. Marit, E. Nursanti, and P. Vitasari, 'Critical Path Method to Accelerate Automotive Maintenance Duration', *International Journal of Scientific & Technology Research*, Vol. 9, No. 03, (2020), pp. 6777-6782.

[32] A. Premono, M. Victor, and H. H.

Sutrisno, 'An experimental study of a car maintenance workshop layout optimization', *AIP Conf Proc*, No. May, (220), pp. 1-8.

[33] N. Siregar, 'Queue System in Automobile Service Sectors', *International Journal of Advanced Science and Technology*, Vol. 29, No. 5, (2020), pp. 13365-13378.

[34] D. Kothandaraman and I. Kandaiyan, 'SS symmetry Analysis of a Heterogeneous Queuing Model with Intermittently Obtainable Servers under a Hybrid', *Symmetry (Basel)*, Vol. 15, (2023), pp. 1-18.

[35] G. Pang, A. Sarantsev, and Y. Suhov, 'Birth and death processes in interactive random environments', *Queueing Syst*, Vol. 102, No. 1, (2022), pp. 269-307.

[36] Y. Chen and W. Whitt, 'Applying optimization theory to study extremal GI/GI/1 transient mean waiting times', *Queueing Syst*, Vol. 101, No. 3, (2022), pp. 197-220.

[37] C. Franco, N. Herazo-Padilla, and J. A. Castañeda, 'A queueing Network approach for capacity planning and patient Scheduling: A case study for the COVID-19 vaccination process in Colombia', *Vaccine*, Vol. 40, No. 49, (2022), pp. 7073-7086.

[38] A. Aghsami, S. R. Abazari, A. Bakhshi, M. A. Yazdani, S. Jolai, and F. Jolai, 'A meta-heuristic optimization for a novel mathematical model for minimizing costs and maximizing donor satisfaction in blood supply chains with finite capacity queueing systems', *Healthcare Analytics*, Vol. 3, (2023), p. 100136.

[39] A. P. Panta, R. P. Ghimire, D. Panthi, and S. R. Pant, 'Optimization of M / M / s / N Queueing Model with Reneging in a Fuzzy Environment', *American Journal of Operations Research*, Vol. 11, (2021), pp. 121-140.

[40] W. Al Okaishi, A. Zaarane, I. Slimani, and I. Atouf, 'A Vehicular Queue Length

Measurement System In Real-Time Based On Ssd Network', *Sciendo*, Vol. 22, No. 1, (2021), pp. 29-37.

[41] A. Aziziankohan, F. Jolai, M. Khalilzadeh, R. Soltani, and R. Tavakkoli-Moghaddam, 'Green supply chain management using the queuing theory to handle congestion and reduce energy consumption and emissions from supply chain transportation fleet', *Journal of Industrial Engineering and Management*, Vol. 10, No. 2, (2017), pp. 213-236.

[42] A. Ltaif, A. Ammar, and L. Khrifch, 'A goal programming approach based on simulation and optimization to serve patients in an external orthopedic department', *Journal of Simulation*, (2022), pp. 1-11.

[43] G. B. Yom-Tov and C. W. Chan, 'Balancing admission control, speedup, and waiting in service systems', *Queueing Syst*, Vol. 97, No. 1, (2021), pp. 163-219.

[44] L. Ravner and J. Wang, 'Estimating customer delay and tardiness sensitivity from periodic queue length observations', *Queueing Syst*, Vol. 103, No. 3, (2023), pp. 241-274.

[45] L. Xiao, S. H. Xu, D. D. Yao, and H. Zhang, 'Optimal staffing for ticket queues', *Queueing Syst*, Vol. 102, No. 1, (2022), pp. 309-351.

[46] W. Dai and J.-Q. Hu, 'Correlated queues with service times depending on inter-arrival times', *Queueing Syst*, Vol. 100, No. 1, (2022), pp. 41-60.

[47] C. Bayliss and J. Panadero, 'Simheuristic and learnheuristic algorithms for the temporary-facility location and queuing problem during population treatment or testing events', *Journal of Simulation*, (2023), pp. 1-20.

[48] S. Lamidi, N. Olaleye, Y. Bankole, A. Obalola, E. Aribike, and I. Adigun, 'Applications of Response Surface Methodology (RSM) in Product Design, Development, and Process Optimization', in *IntechOpen*, (2022).

[49] P. Kemper, D. Müller, and A. Thümmler, 'Combining response surface methodology with numerical models for optimization of class-based queueing systems', in *Proceedings of the International Conference on Dependable Systems and Networks*, (2005), pp. 550-559.

[50] S. Ahmed and F. Amagoh, 'Application of Simulation and Response Surface to Optimize Hospital Resources', *International Journal of Mathematical and Computational Sciences*, Vol. 4, No. 6, (2010).

[51] X. yue Xu, J. Liu, H. ying Li, and J. Q. Hu, 'Analysis of subway station capacity with the use of queueing theory', *Transp Res Part C Emerg Technol*, Vol. 38, (2014), pp. 28-43.

[52] M. Baghery, H. P. Abgarmi, S. Yousef, A. Alizadeh, and H. Mahmoudzadeh, 'Simulation-Based Optimization for Improving Hospital Performance', *International Journal of Hospital Research*, Vol. 6, No. 2, (2017), pp. 72-85.

[53] S. Asadzadeh, B. Akhavan, and B. Akhavan, 'Multi-objective optimization of Gas Station performance using response surface methodology', *International Journal of Quality and Reliability Management*, Vol. 38, No. 2, (2021), pp. 465-483.

[54] A. R. Gottu Mukkula and R. Paulen, 'Robust design of optimal experiments considering consecutive re-designs', *IFAC-PapersOnLine*, Vol. 55, No. 7, (2022), pp. 13-18.

[55] J. Antony, E. Viles, A. F. Torres, T. I. de Paula, M. M. Fernandes, and E. A. Cudney, 'Design of experiments in the service industry: a critical literature review and future research directions', *The TQM Journal*, Vol. 32, No. 6, (2020), pp. 1159-1175.

[56] W. Sugianto, R. Haq, and A. Haq, 'Analysis of Factors that Influence Automobile Workshop Queue Performance Using Design of Experiments', *Journal of Optimization in Industrial Engineering*, Vol. 16, No. 2, (2023), pp. 303-314.

[57] N. R. Draper and J. A. John, 'Response-surface designs for quantitative and qualitative variables', *Technometrics*, Vol. 30, No. 4, (1988), pp. 423-428.

[58] M. R. Galankashi, E. Fallahiarezoudar, A. Moazzami, N. M. Yusof, and S. A. Helmi, 'Performance evaluation of a petrol station queuing system: A simulation-based design of experiments study', *Advances in Engineering Software*, Vol. 92, (2016), pp. 15-26.