SIMULATION OF ROUTING POLICIES FOR A MULTI-CLASS DOCKING SYSTEM OF AN AUTOMOBILE MANUFACTURER

by

Murat Cesur

B.S., Industrial Engineering, Middle East Technical University, 2013

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering

Boğaziçi University

2017

ACKNOWLEDGEMENT

First of all, I would like to express my special appreciation and thanks to my advisor Assoc. Prof. Aybek Korugan for his guidance, continuous support, motivation and immense knowledge.

My thanks and appreciation goes to my examining committee members Prof. Zeynep Akşin and Assoc.Prof. Gönenç Yücel for their valuable suggestions and comments.

I truly appreciate the support of my company for my graduate study. I would also like express my special thanks to Mehmet Karaca, Altan Aytaç, Onur Yalçın and S.Nazlı Eken from Tofas Türk Otomobil Fab.A.S. for their support.

I gratefully recognize the support of TÜBİTAK – Turkish Technological and Scientific Research Institute - for providing a scholarship during my graduate education with the program of 2210/A-National Scholarship Programme for M.Sc. Students (within the scope of 2211-National Scholarship Programmes for Graduate Students)

Most importantly, I would like to express my gratitude to my father and mother, Sabri and Gülfiye, who encouraged and helped me at every stage of my personal and academic life. I thank for their support, patience and love. They deserve to have their names on the diploma more than I do.

ABSTRACT

SIMULATION OF ROUTING POLICIES FOR A MULTI-CLASS DOCKING SYSTEM OF AN AUTOMOBILE MANUFACTURER

Supply chain operations are the vital functions of today's huge and complex organizations. Although most of these activities are inevitable due to the different needs of customers, eliminating some non-value-adding activities provides significant cost reduction opportunities. Considering internal logistics activities, it is seen that material-receiving process includes many types of inefficiencies. To illustrate; waiting trucks, unloading materials at unsuitable docks or huge distances from docks to the warehouses are the inefficiencies resulting in extra costs without adding value for the product. The presented study is based on truck unloading and material acceptance processes in an automotive plant. After entering the factory, trucks wait in queues if all unloading docks are busy in that time and then they are directed to suitable docks. "What kind of a queuing system should be designed?", "Which truck should be selected from the queue when an unloading dock becomes idle? or "Which unloading dock should be selected if there are some idle docks?" are the main decisions to be made in daily operations. This study aims to investigate different truck routing and dock selection policies supporting the decision making process. By using ARENA simulation software, studied policies are simulated in various environments with different input parameters. Resulting average waiting time per truck and average covered distance to carry unloaded materials from unloading docks to warehouses are investigated as main performance measures of the system. Due to the multi-objective setting of the study, pareto analysis is utilized by using two-dimensional trade-off curves. Analyzing system outputs, some generalized conclusions about the effectiveness of the proposed routing policies are drawn. Then, an actual data set is obtained from an automotive factory and a real life case study is investigated. After simulating the process with different routing policies, the efficient ones are determined for the automotive factory.

ÖZET

BİR OTOMOBİL ÜRETİCİSİNDE ÇOK SINIFLI RAMPA SİSTEMİ İÇİN ROTALAMA POLİTİKALARININ SİMÜLASYONU

Günümüzdeki büyük ve oldukça karmaşık organizasyonlarda tedarik zinciri operasyonları hayati bir yere sahiptir. Müşteri ihityaçlarını karşılayabilmek için bu aktivitelerin bir çoğunun kaçınılmaz olmasına karşın, katma deter sağlamayan kayıpların elenmesi ciddi miktarda maliyet imdirimi sağlayabilmektedir. İç lojistik aktiviteleri düşünüldüğünde, malzeme alımı/kabulü bir çok kaybı içinde barındırmaktadır. Örneğin, gelen tırların boşaltılmak için beklemesi, malzemelerin uygunsuz kabul noktalarında boşaltılması veya ambarlar ile malzemenin kabul edildiği noktalar arasında oldukça uzun mesafelerin katedilmesi; katma değer sağlamaksızın ek maliyer yaratmaktadır. Bu çalışmada, bir otomobil fabrikası baz alınarak, fabrika içerisinde gerçekleşen tır boşaltma ve malzeme kabul işlemleri incelenmektedir. Bu süreçte, fabrika içerisine giren tırlar, eğer o anda uygun bir indirme noktası boş değilse sıraya girerek beklemekte ve sonrasında boşaltma noktaları uygun olduğunda uygun rampalara yönlendirilmektedirler. “Ne tip bir kuyruk sistemi tasarlanmalı?”, “Bir boşatma noktası boşaldığında, bekleyen tırlardan hangisi seçilmeli?” veya “Birden çok indirme noktasının boş olması durumnda, gelen tır hangisine yönlendirilmeli?” soruları günlük operasyonlar sırasında verilmesi gereken kararlardır. Bu çalışmada, sistem performansını iyileştirebilmek amacıyla, çeşitli tır yönlendirme ve indirme noktası seçim algoritmaları çalışılmıştır. Oluşturulan bu algoritmalar farklı girdi ve sistem kurguları içeren senaryolarda ARENA simulasyon programı ile simule edilmiştir. Tüm bu kararların sonucu olarak; tırların bekleme zamanları ve indirme noktaları ile ambarlar arasında malzemelerin katettiği uzaklık gibi performans ölçütleri ortaya çıkmaktadır. Sistemin birden çok performans ölçütü içeren yapısından dolayı, ödünleşme eğrileri kullanılarak pareto analizi yöntemi ile sonuçlar incelenmektedir. Çalışmanın son aşamasında, bir otomobil fabrikasında gerçekleşen malzeme kabul operasyonları önerilen farklı algoritmalar ile simule edilmiş ve bu fabrika için en etkin algoritmalar önerilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

d	Dock type
h	Half-length in confidence interval analysis
Lq	Queue length
n	Replication number
$p_{t,d}$	Penalty value of unloading a type *t* truck at unloading dock *d*
r	Relative precision level
$sp_{t,d}$	Static probability of assigning truck t to dock *d*
t	Truck type
Wq	Waiting time in the queue
$x_{t,d}$	Ratio of trucks from truck type *t* assigned to unloading dock *d*
α	Significance level
λ_t	Arrival rate of type *t* trucks
$\mu_{t,d}$	Unloading rate of dock *d* for type *t* trucks
ρ	Traffic intensity rate

LIST OF ACRONYMS / ABBREVIATIONS

Created	Number truck created during the simulation
DAVG	Average of a time-persistent variable
FIFO	First in first out
FSF	Fastest servers first
GDP	Gross domestic product
K-S	Kolmogorov-Smirnov
LISF	Longest idle server first
MDP	Markov decision process
Min_CbPr	Minimum combined parameter policy
Min_P	Minimum penalty policy
Min_UT	Minimum unloading time policy
NR	0/1: Resource is busy(1) or idle (0)
NQ	Number of trucks in queue
QIR	Queue-and-idleness-ratio
Rand	Randomized routing policy
SCM	Supply chain management
Serviced	Number of trucks unloaded
SPT	Shortest process time
ss	Steady state
StAllc	Static allocation policy
StAllc_ARAdms	Static allocation policy with attribute dependent admission under assignment rate constraint
StAllc_THAdms	Static allocation policy with attribute dependent admission under threshold queue length constraint
TAVG	Average of a tally variable
TH	Threshold
WH	Warehouse

1. INTRODUCTION

Considering the cost and quality driven competitive world of today, it is seen that logistics activities have critical economic importance for the companies. Since logistics related activities create considerable amount of cost, they are seen as cost centers of the organizations. Therefore, any improvement in logistics costs provide important competitive advantage for the companies in providing a leverage for reducing the cost of their products or services. A closer look at the logistics cost breakdown shows that the transportation and material procurement/storage operations are the most costly operations as seen in Figure 1.1 [1].

Although, cost of transportation activities are inevitable in many cases, elimination of losses due to inefficient implementations can be possible. To illustrate, acceptance of incoming materials is a key operation linking transportation and warehousing activities. Due to the complexity of dock planning, material unloading and warehousing operations, dock selection for material unloading and routing trucks to these selected docks are done based on the intuition of operators in many organizations. Therefore, waiting time of trucks to unload materials to the docks or unloading the materials at improper docks result in significant inefficiencies.

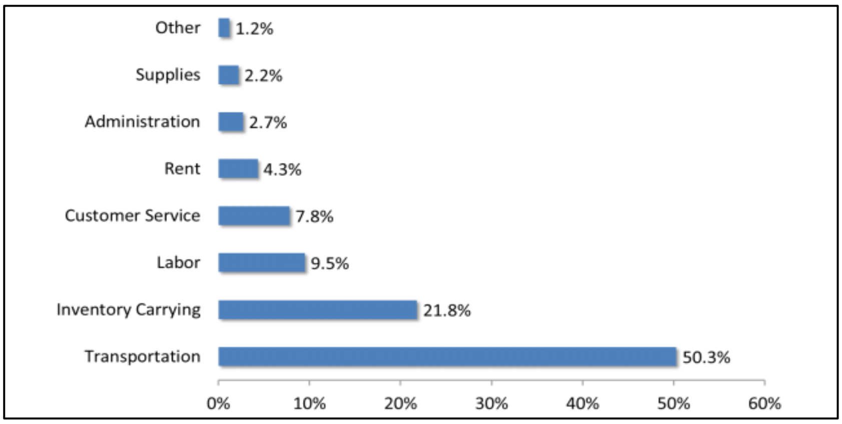

Figure 1.1. Logistics costs breakdown

Automotive is one of the uppermost sectors including high level of material movement both internally and externally. An average car is produced by assembling more than 1,000 parts and more than 10,000 part numbers are managed in an average automotive factory.

Therefore, plenty of external transportation routes and internal handling movements take place during daily logistics activities.

This study is based on a problem encountered during the incoming material unloading/acceptance processes in an automotive factory. The investigated process starts with the arrival of the trucks to the automotive factory. Supplied materials are received by trucks and unloaded at the gates of the warehouses (unloading docks). Trucks arrive continuously to the factory in a given day and they are classified based on the materials they carry. Some of these trucks carry domestic materials that are procured from geographically close suppliers while some of them carry imported materials from remote suppliers that have longer lead-times. Depending on the material types they carry, the arrival rates of trucks the vary and the inter-arrival times of different truck types are independent from each other. Moreover, due to the traffic intensity, weather conditions, supplier problems or transportation problems, arrival times show a considerable level of stochasticity.

After the arrivals of trucks, materials are unloaded at unloading docks and carried to the warehouses where they are stored. Multiple warehouse areas are located inside the factory according to the need of the production line and each special warehouse is used to stock one type of material. Unloading docks, on the other hand, are the gates of the internal warehouses where arriving materials are accepted and unloaded. Afterwards, accepted materials are carried to their respective warehouses. Considering the flow of a special type of incoming material from the unloading dock to its warehouse, the dock selection for material unloading is important since unloading at the farther docks results in extra material handling activities. On the other hand, waiting for the nearest ideal dock to unload materials may increase waiting time of the truck if its ideal dock is busy at the time when the truck arrives. Thus, unloading docks are heterogeneous (non-identical) for different types of trucks. Both unloading times and material movement distances from unloading docks to the warehouse areas vary based on the selection of unloading docks. Arrival rates of trucks, number of unloading docks, unloading times in different docks and distance between docks and storage areas are taken as input parameters of the system.

Considering the mentioned system setup, selecting and implementing an effective truck routing policy is critical to improve efficiency of the system. A routing policy makes three main decisions. Firstly, it decides the queuing system setting. Single queue or dock-dedicated queues settings can be selected to implement. Secondly, if dock dedicated queues are constructed, allocation decision for an arriving truck is made to select one of the dock-dedicated queues. Lastly, when a dock completes its operation, truck admission decision is made according to the policy algorithm and one of the waiting trucks is called for the unloading operation.

In such a system setting, this study aims to investigate implementable and efficient routing policies for truck unloading operations with the presence of multi type heterogeneous trucks and multi type heterogeneous unloading docks. Different policies are investigated in different system settings and simulation results are analyzed. The remainder of the study is organized as follows:

Related literature review is presented in Chapter 2 and problem definition and aim of the study is given in Chapter 3. Problem structure and appropriate routing policies are modelled in Chapter 4. Design of simulation experiments and scenario analysis are explained in Chapter 5. With different system parameters, several experimental scenarios are constructed in this chapter. Sensitivity analysis, results, trade-off curves and efficient frontiers are investigated and the results are presented. Chapter 6 contains a real life case study from an automotive factory. After statistical input analysis, different policies are simulated and results of different policies are compared. Main conclusions and possible future research areas are discussed in the final chapter.

2. LITERATURE REVIEW

In this chapter, related literature is presented and contents of the past studies are introduced briefly. This thesis is mainly focused on routing policies to guide arriving trucks to the unloading docks. Due to the limited capacity of docks, material unloading operations cannot be started as soon as a truck arrives. In such an environment, arrived trucks wait for the busy docks to complete their operations. Hence, deciding about the queuing system settings and using queuing theory results to investigate for efficient implementations are relevant study areas.

Queuing theory implementations in service and manufacturing systems attract many researchers to investigate the different environment settings and develop new solution procedures. Considering the past studies, starting from single type of customer and server with single queue settings, considerations are developed to heterogeneous customers and heterogeneous servers settings with multiple queues.

The aim of this thesis is to investigate an efficient and implementable truck routing algorithm with an efficient queuing system setting. In the problem of interest, there are multiple types of trucks and multiple types of unloading docks. Each truck type can be unloaded at a specific type of dock but the distance between the selected dock and warehouse and unloading time are varied for each truck type-dock type pair.

As the environment becomes such a complex structure with heterogeneous customers and heterogeneous servers, obtaining exact solution methodologies get harder as seen in literature. Under high level of uncertainty of arrival and service rates, simulation studies and numerical analysis become popular. Since real life problems occur in very dynamic environments, the proposed policies should be easy to implement. Therefore, easily applicable routing policies and customer selection rules are very useful and preferable in complicated environments in many implementation areas.

Queuing theory and routing policy implementations are the combined methodologies used in many different disciplines. Since most of the real life implementations comprise of very complex processes, using exact analysis is very challenging. Due to this fact, simulation studies are common in the literature. Implementation of the routing policies with heterogeneous servers are common in service sectors like call centers or hospitals. Yet they do not exist or are rarely applied to logistics systems in manufacturing sector. Considering the aim of the study and main characteristics of interested system settings, literature review is presented in three parts.

2.1. Queuing Systems and Routing Policies

In accordance with the aim of this thesis, papers related to queuing systems and routing rules are investigated in this part of the literature review. Especially, studies with multi classes of servers/customers and multi-dimensional performance measures are reviewed. Considering the traditional queuing systems and routing policies literature, it is seen that waiting time of customers is the most widely used performance measure. However, Vericourt & Zhou [3] asserts that compatibility of server and customer is also an important requirement for the environments where servers or customers are not identical to each other. In addition to the waiting time of customers, they also concentrate on resolution probability of service. Resolution probability is defined as the probability that selected server solves the assigned customer's problem. In their paper, a call center system with heterogeneous servers and homogenous customers with retrial characteristics is investigated. Retrial characteristics means that; if the problem of a served customer is not cleared, the customer comes to the system again. They determine the optimal policy based on an MDP model of the problem. Then, they compare sub-optimal policies numerically. In the paper, a routing policy that ranks the servers according to their $p.\mu$ rates is proposed. Here, μ is the expected service rate that is the direct measure of the server speed and p is the resolution probability expressing the compatibility of server and customer. Customer prefers the available server with the highest $p.\mu$ value to get service. They show that if only p or only μ values are diversified, then $p\mu$ rule is optimal. They also propose "$p\mu$-t rule" where the customer is always send to the server with higher $p.\mu$ rate until the queue of the server reaches to the threshold value. It is also explained that the optimal routing policy for the system with two type of servers is characterized by using "$p\mu$-t rule".

Armony [4] also describes an environment where heterogeneous servers handle homogenous customers waiting in a single queue. Fastest Servers First (FSF) rule is proposed considering the heavy traffic and heterogeneous server regimes. Although FSF rule is a purely μ-based rule she also asserts that her methodology can be extended to use pμ rule suggested in Vericourt *et al.* [3] where multiple heterogeneous servers have different service times (μ) and clearance rates (p). Although FSF rule directs the customers to the faster servers and helps to minimize both service and waiting times of customers, it can lead to very different server utilization rates. In other words, faster servers can be over utilized compared to slower servers in order to reduce the service times of customers.

To balance the workload of different servers with heterogeneous servers and homogenous customers setting, Atar [7] proposes a routing policy that directs the coming customer to the idle server that has been idle for the longest time. This policy is stated as Longest-Idle Server First (LISF) policy. Since this policy mainly focuses on eliminating unfair workloads of servers, customer-oriented performance measures like waiting time of customers or server selection rate are disregarded.

Armony and Ward [8] study to balance the tradeoff between minimizing waiting time and maximizing the fairness during the assignment of the workload to the servers. They formulate an optimization problem that minimizes expected customer waiting time subject to a workload balancing constraint. Since the solution structure is not simple to implement, they propose a routing rule that combines Longest-Idle Server First (LISF) policy and Fastest Servers First (FSF) policy logics. The proposed policy calculates a threshold value for the number of customers in the system and assigns the customers to the faster servers if the number of customers is higher than the threshold value. If the threshold value is not reached, assignment of the customers are made according to LISF policy.

Gurvich and Whitt [9] study multi type heterogeneous customer and heterogeneous server systems. They consider a group of policies called Queue-and-Idleness-Ratio (QIR) rules. The rule concerns about both the queue length of different types of customer and idleness rates of servers. When a server becomes idle, the customer having queue length exceeding the predetermined proportion of the total queue length is selected. When a customer comes to the system, on the other hand, the server whose idleness exceeds the predetermined proportion of the total idleness most is selected. They state that Queue-and-Idleness-Ratio (QIR) rules results in an integrated output that considers both waiting time of customers and idleness rates of servers.

Zhan and Ward [10] propose a routing control system that gives static priority to some servers and uses dynamically determined threshold values to give state-based priority to other ones. Their methodology considers both service speed and resolution probability of a server. By using simulation study, they constitute efficient frontier graphs regarding waiting time and customer satisfaction. Their study includes only one type of homogenous customer whereas servers differ from each other with respect to both service rate and service quality.

2.2. Queuing Systems Simulation Studies

As discussed earlier, it is either very hard or impossible to find an exact solution for the complex queuing and routing systems. Therefore, for the problem of interest in this thesis, a simulation-based approach is more suitable for the analysis. To this end, in this section we review simulation studies and numerical analyses used to find efficient routing and queueing algorithms for complex environments.

In their paper, Tits *et al.* [13] investigate whether average queue length and average waiting time formulae for exponential arrival and waiting time distributions stated in the paper of Kleinrock *et al.* [14] are also valid for non-exponential distributions by using simulation. This study was one of the initial simulation studies that makes comparison between simulation outcomes and analytical results.

Besides their study, many papers use simulation models to understand the characteristics of complex systems. Glynn *et al.* [15] wrote about simulation methods for queues as an overview study. In this paper, they give an overview of simulation methodology for queueing systems. Input, output analysis and variance reduction methods are covered in the study. Janng *et al.* [16] study about the estimation of the mean waiting time of a customer subject to balking with a simulation study. This model considers general inter-arrival and service time distributions and uses simulation and regression methods to estimate mean waiting times and effect of the factors on the average waiting time.

Mehrotra *et al.* [11] try to characterize the efficient routing policies in the presence of heterogeneous customers and heterogeneous servers. Like Zhan and Ward [10], they consider waiting time and resolution probability as performance measures. In this paper, many rules are proposed for different working conditions and these rules are classified as minimum waiting time oriented and maximum resolution rate oriented rules. Hybrid rules, like p.μ rule and threshold rules that balance the two extreme performance directions are also suggested. After routing policy characterization, a call center is simulated and efficient frontiers are constructed. All these results are obtained by statistical output analysis with the result of the simulation studies.

2.3. Warehouse-Unloading Truck Queue Simulation Studies

Considering queuing and routing policies and their simulation implementations, it can be seen that queueing simulations are widely used in many areas from service systems to manufacturing systems. Similarly, logistics operations are important areas in which queuing theory and simulation methodologies are used effectively. Although analytical results obtained by queuing theory (as discussed in Chapter 2.1) are used in logistics systems where trucks, unloading areas or unloading times are homogenous; for the heterogeneous or non-stationary systems simulation and numerical studies are extensive. In this thesis, truck unloading operations will be simulated by using queuing theory and simulation with the integration of dock optimization policy decisions. Considering similar topics in the literature the following studies can be referred:

In Woottichaiwat's study [18], a queuing system of supply trucks in freight unloading process in a private port in Songkhla province is investigated. Different classes of trucks are coming for unloading and each truck has different operation time. With real data collection and system process investigation, simulation model is constructed. Different truck selection algorithms (FIFO, SPT, …) and several possible process paths are simulated and output analysis are performed. As a result, a new system is proposed with waiting time reduction and server utilization improvement.

Gopakumar *et al.* [19] use a queueing simulation approach for dock allocation in a food distribution center in their study. The study mainly focuses on the product reception process of a warehouse. Discrete event simulation is conducted in ARENA software and inbound activity time distributions are given as an input. Different process paths and dock allocation scenarios are studied with the constructed simulation model. Ultimately, results are discussed and some policies are selected to implement.

Iannoni *et al.* [20] investigate the truck waiting times in the queues of the plant reception area in a sugarcane plant. They analyze the performance of different configuration and policies for receiving operations. Discrete simulation techniques are applied to study the reception area processes. With several sensitivity analysis, different effects are examined and input/output analysis are performed to reach an efficient receiving process for the plant.

Liong *et al.* [21] construct a simulation methodology with ARENA software to find a strategy that optimizes the residence time of any truck in the warehouse. They consider all warehousing operations from loading to unloading operations in their study. In this paper, four improvement models are developed and compared. They also focus on the basics of the simulation study with ARENA software and show the effects of the level of inputs on the performance measures by implementing sensitivity analysis.

Results obtained from literature review can be summarized as follows:

- Especially less complex processes with homogenous-type customers/servers are modelled and many important exact analytic formulae for performance measures are constructed.

- However, it is very difficult to get the analytic solutions or characterize the exact optimal routing policies for the performance measures of the queue systems with complex scenarios especially with heterogeneous multi class customer and server cases. In such cases numerical methods, simulation studies and regression methodologies are used widely.

- Although warehouse movements like truck unloading or loading operations are quite tangible and suitable processes to model with queuing systems and routing policies, integrated studies in this area is not very extensive. Especially, integrated studies including truck routing and dock selection policies for multi truck classes and nonhomogenous dock types are very rare. Considering logistics systems literature it is also seen that two different cost parameters (penalty and waiting time) and multi-objective system settings are missing.

In this thesis, queuing and routing policies are investigated in heterogeneous unloading docks and heterogeneous truck types setting. Waiting times of trucks and penalty rates representing internal handling costs of materials are considered together. Easy and implementable routing and prioritization policies taking into account both of these costs are investigated through the study.

Most of the manufacturing companies suffer from long waiting times of truck inside the factory and resulting congestion due to the lack of an implementable policy. Hence, the results of this study will contribute to improve logistics processes. Additionally, holistic approach that considers both external logistics costs (truck waiting costs) and internal handling costs (penalty rates representing carrying materials from dock to the warehouse) will create opportunities to increase the efficiency for many manufacturing plants.

3. PROBLEM DEFINITION AND AIM OF THE STUDY

The problem investigated in this thesis is inspired by a real life problem in an automotive factory. The system involves three unloading docks and three types of trucks. Trucks are classified in three groups according to the material types carried as “Domestic Material Trucks”, “Import Material Trucks” and “JIS Material Trucks”. These trucks are arriving to the plant with different arrival rates. Due to the internal and external environmental effects, the inter-arrival times are random and independent of each other. Currently arriving trucks are routed to unloading docks with a random policy and these trucks enter to the queue of one of the unloading docks where they wait in dock dedicated FCFS queues to unload their materials. Unloading time for each dock type-truck type pair has a different expected unloading time with possible different distributions.

Another important performance measure is the distance between the unloading dock where the truck is routed and the warehouse area of the unloaded material. Each material type, that determines the type of the truck, has a dedicated stocking area inside the plant. After the unloading operation, materials are carried these areas. The distance between the selected dock and warehouse area of unloaded material results in inefficiencies in the system. Carrying materials for long distances causes labor cost due to the material handling. Additionally, high level of traffic congestion in warehouse areas blocks the logistics operations. Therefore, there is a distance matrix for each unloading dock-truck type pair. The schematic representation of the system is given in Figure 3.1

In addition to obtaining an efficient solution to the specific problem explained above, constructing a solution methodology and generalizing the results for the similar classes of problems is one of the main objectives of this thesis. It is also intended to reach generalized results that can be used in many areas in which the core of the problem definition is based upon the problem of interest in this study.

Due to the stochasticity of arrival times of trucks and unloading times at docks according to the truck types, the complexity of the problem is quite high. Building an analogy between the stated problems and queuing theory, there are multi-type heterogeneous customers (trucks) and multi-type heterogeneous servers (unloading docks) in the system. Actually, this study aims to investigate implementable and efficient routing policies for truck unloading operations with the presence of multiple type heterogeneous trucks and multiple type heterogeneous unloading docks.

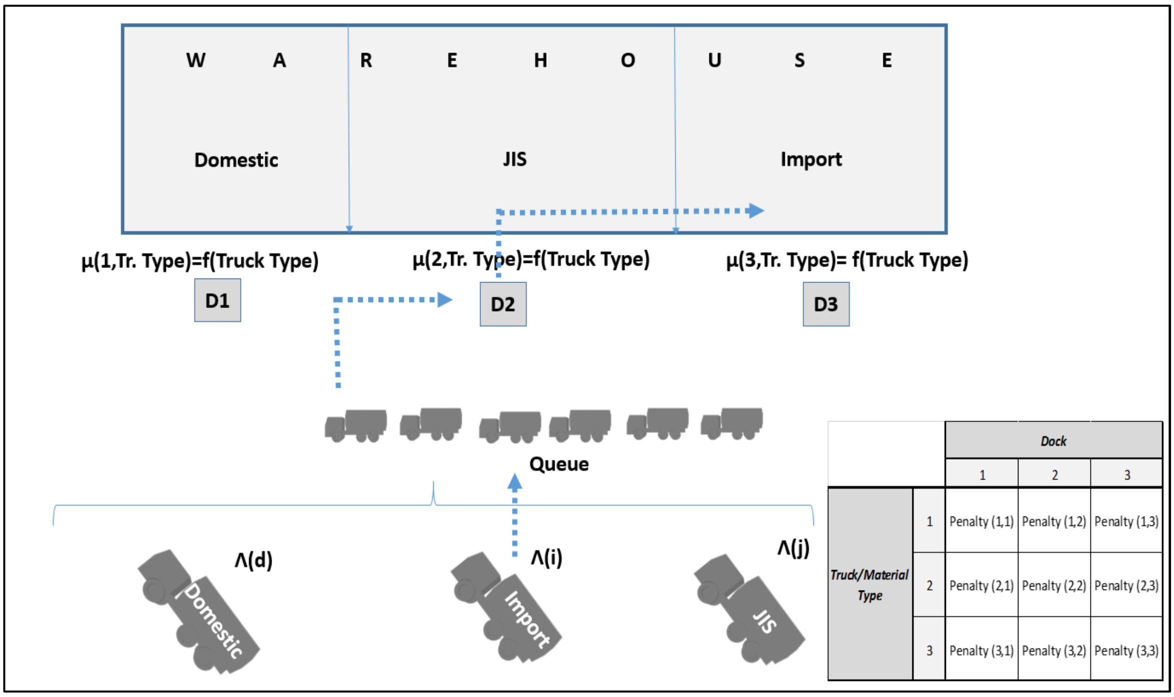

Figure 3.1. Schematic representation of the system

To see the results of implemented policies in different environments, studied methodologies are simulated with different input parameter rates. In order to understand the effects of uncontrollable factors like traffic density, penalty and service time correlations or dispersion of penalty values for specific docks, experimental simulation design studies are conducted. Observing the outputs of investigated policies in different experimental environments and constructing the tradeoff curves that shows the tradeoffs between two performance measures (waiting time vs. penalty rate –distance-) is one of the most important goals of this study. Since metrics of performance measures (distance and waiting time) are different from each other, these performance measures are investigated separately instead of combining them into one measure. As a result, trade-off curves are used to conduct pareto analysis in the cases when waiting time and distance minimization objectives conflict.

4. MODEL DEVELOPMENT

After constructing representative models for all routing policies developed, ARENA and GAMS software packages are used to run the models. GAMS is used to solve linear optimization models while the simulation models are developed in ARENA Simulation software. Afterwards, different policies are simulated in several environments. Input Analyzer and Output Analyzer modules of ARENA is benefited to construct input distributions and make statistical analysis with the outputs of the simulation models. In this chapter, general model, assumptions and simulated models with different routing policies are explained in detail.

4.1. Model and Assumptions

The model framework setting is constructed based on the presence of multiple truck types (indexed by *t*) and multiple dock types (indexed by *d*). Different types of trucks come into the system to empty their loads. These trucks can form a single line or multiple lines for each dock according to the predefined policy.

Arrival rate of type *t* trucks is shown by λ_t. Each truck type is unloaded with different service rate at each different dock. The unloading rate of truck *t* at dock *d* is shown by $\mu_{t,d}$. Considering the dock preferences of different truck types, penalty cost of $p_{t,d}$ occurs if a *t* type of truck is unloaded at dock d. As explained before, the penalty represents the extra handling cost due to the distance from selected unloading dock to the warehouse of the unloaded material type.

Each arriving truck acts according to the implemented routing policy. The routing policy determines which truck should be directed to which unloading dock at each decision epoch. It is assumed that the system (dispatchers) have full visibility of parameters during the policy implementation phase. Average waiting time and average penalty rate per truck are seen as the two important performance measures of the system. All other input parameters like arrival, service and penalty rates are assumed to be stationary and do not change with time. Additionally, for the simulation studies inputs are determined such that the system reaches to steady state and all analysis is made with steady state values.

4.2. Routing and Queuing Policies

Different routing policies are investigated to select an efficient policy for the problem of assigning trucks to unloading docks. Waiting times of trucks and average penalty rate occurred due to truck-dock matchings are two important performance measures of the system. Due to the presence of two different performance measures with different metrics, both of them are obtained separately and integrated output analysis are performed by using 2-dimensional trade-off curves. Due to the multi-objective setting of the system, some of these policies are powerful for waiting time minimization while others are effective for penalty minimization aspect. To balance between these two extreme edges, some hybrid policies are also studied. The brief explanations of these policies are mentioned in the following sections.

4.2.1. Single Queue Policies

In single queue policies, only one queue is constituted and all arrived trucks are waiting in the same queue. Unloading truck is selected from this single queue for all unloading docks.

<u>4.2.1.1. Minimum Unloading Time Policy (Min_UT).</u> One of the simplest policy that aims to minimize customer waiting time is Fastest Servers First rule as stated in Armony (2005). The aim of the policy is assigning trucks to the available docks that operate faster for that type of trucks. Although this policy is expected to be very effective to minimize total waiting time of trucks, it totally ignores large penalty values. Considering the decision points, the algorithm of minimum unloading time policy can be stated as follows:

- An unloading operation is completed and at least one unloading dock is empty.

  (i) There is no truck waiting for unloading: There is no action to take. Available docks wait for arrival of at least one truck.

(ii) At least one truck is in the queue waiting for unloading: Select the "waiting truck-available dock" pair that has the minimum unloading time. If there is a tie, select one of these candidate pairs randomly.

- A Truck comes to the queue for unloading.

(i) There is no available dock for unloading: There is no action to take. The arriving truck enters to the queue and waits for unloading.

(ii) At least one dock is available for the unloading operation: Select the "waiting truck-available dock" pair that has the minimum unloading time. If there is a tie, select one of these candidates randomly.

4.2.1.2. Minimum Penalty Policy (Min_P). The same logic used in minimum unloading time policy is applied for the implementation of minimum penalty policy. The only difference is that penalty values is used as selection criterion instead of unloading time values. Although this policy is expected to be very effective to minimize total penalty, it ignores possible long waiting times of trucks due to dock selection decisions.

4.2.1.3. Minimum Combined Parameter Policy (Min_CbPr). Instead of using pure unloading time or penalty criterion, a combined parameter is created by considering unloading time and penalty rates. Then routing decisions are made based on this created parameter in place of pure unloading time parameter in minimum unloading time rule.

For each (t,d) – (truck type, dock) pair, unloading time and penalty rate are predefined parameters. In order to create a common parameter, both unloading time and penalty values are scaled from 1 to 10. Minimum of unloading time (penalty) value is scored as 1 and maximum of unloading time (penalty) value is scored as 10. Then the intermediate values are scored from 1 to 10 according to their position in reference to minimum and maximum values. By multiplying the unloading time and penalty rate scores, combined parameter is constructed for each (t,d) pair.

Min_CbPr policy is a single queue hybrid policy that considers both service times and penalty rates at the same time. Since metrics scale of these two parameters are different, scaling process is implemented. Afterwards, multiplication of these two-scaled parameters highlights the dock-truck matchings that have fair service time and penalty rates. In other words, truck-dock pairs having fair service time and penalty rate are preferred to the pairs having excessive service time (penalty rate) and very low penalty rate (service rate). The aim of using multiplication operation for two different parameters is not to undervalue one of these parameter for the sake of the other but highlight truck-dock pairs that have acceptable levels for both parameters (penalty rate and waiting time).

4.2.2. Dock Dedicated Queue Policies

In dock dedicated queue policies, different queues are constituted for each unloading dock separately. Since there are more than one queue, arriving trucks are allocated to one of these dock-dedicated queues when they arrive. After the first allocation, some of the dock dedicated queue policies allow truck admission from different docks' queues, while others force the allocated trucks to wait in their own dock dedicated queues until completion of their unloading operations.

4.2.2.1. Randomized Routing Policy (Rand). Randomized routing rule represents the scenario where there is no clever algorithm to route the arriving trucks to unloading docks. Each coming truck is directed to one of the unloading dock queue randomly. The results of this scenario can be used as base reference outputs of the defined system. Randomizing routing rule assigns the arriving trucks to the dock dedicated queues with equal probabilities and a FCFS admission policy is used for each dock-dedicated queue. Implementing this policy is equivalent to construct different M/M/1 queues for each dock.

4.2.2.2. Static Allocation Policy (St_Allc). In order to minimize expected average penalty rate per truck, a deterministic optimization model can be used as explained in the study of Mehrotra *et al.* [11]. This optimization problem is solved once before the simulation study to find the ideal assignments rates of arriving trucks to the unloading docks. Therefore, $x_{t,d}$ values (assignment rates) that represent the probability that truck *t* is directed to unloading dock *d* are predetermined. Since unloading capacity of each dock is limited, in some cases, tuck type arrivals can be assigned to more than one dock with certain probabilities. The $x_{t,d}$ values can be determined by solving following optimization model:

Min $\sum\sum_{t,d} p_{t,d} * x_{t,d} * \lambda_t$ (4.1a)

S.t

$\sum_d x_{t,d} = 1$ for t ε *{1,2,3}* (4.1b)

$\sum_t x_{t,d} * \lambda_t * (\frac{1}{\mu_{t,d}}) < 0.85$ for d ε *{1,2,3}* (4.1c)

$x_{t,d} \geq 0$ for t ε *{1,2,3} and* d ε *{1,2,3}* (4.1d)

Objective function of the model (4.1a) minimizes total expected penalty by deciding assignment rates ($x_{t,d}$). Constraint (4.1b) assures that all trucks of type *t* are assigned to one dock. Unloading capacity constraint for each dock is satisfied via equation (4.1c). Non-negativity constraints for assignment rates are provided by equation (4.1d).

By solving the optimization model, assignment rates $x_{t,d}$ of type *t* truck to the unloading dock *d* is determined to minimize expected total penalty of the system. Then, during the simulation phase, each arrival truck from type *t* is randomly assigned to dock *d* queue with the probability of $x_{t,d}$. By assigning arriving trucks to the docks according to $x_{t,d}$ values, a dedicated queue is constructed for each unloading dock according to predetermined assignment rates. Actually, the setting of this rule is the same as the randomized rule except for the selection prioritization of some queues according to the assignment rates determined in the optimization model.

To constitute a stable system, optimization model is run with constraint (3) that limit the average utilization level of each dock by 0.85. When the maximum utilization level for a dock is bounded as less than 1 (<1), the utilization rates of low-penalty docks are excessively high in this policy. Due to very high level of utilizations, system cannot reach steady state in reasonable time. In order to eliminate this problem, 0.85 (instead of 1) is used as a limit for average utilization level. By this way, maximum expected utilization of a dock is assured to be less than 0.85. Therefore, some amount of free time is set for highly utilized servers to tolerate the stochasticity of the system and steady state is obtained in reasonable time intervals for traffic intense scenarios.

4.2.2.3. Static Allocation Policy With Attribute Dependent Admission Under Assignment Rate Constraint (StAllc_ARAdms). Static allocation policy focuses on minimizing expected penalty rates and ignores waiting time minimization of trucks. Since trucks are assigned to the dock dedicated queues statically at the beginning of the process, making unloading operation at a different idle dock cannot be performed to reduce waiting time of trucks. This policy forces the assigned trucks to wait in the dedicated dock queue while some of other unloading docks are idle. Therefore, long waiting times occur for the sake of minimum penalty rates. Due to these characteristics, Static allocation policy provides a good lower bound reference for penalty rates but it is not preferable to implement in many real life environments due to excessive waiting times.

To eliminate the mentioned disadvantage, a hybrid policy is used to unify minimum combined parameter policy and Static allocation policy. This hybrid policy is implemented similar to minimum combined parameter policy but type *t* truck and *d* unloading dock pair is marked as incompatible if the optimization model gives $x_{t,d}$=0. Hence, the rule does not allow the admission of truck *t* by dock *d* if the $x_{t,d}$ value is zero. The policy algorithm can be summarized as follows:

- An unloading operation is completed at dock *d*.

  (i) There is no truck waiting for unloading at dock *d* queue and all "waiting truck-available dock" pairs in different queues have values of $x_{t,d}$=0 : There is no action to take. Available docks wait for truck arrival.

  (ii) There is no waiting truck in the queue of dock *d* and at least one truck from the other queues satisfies $x_{t,d}$>0: Investigate all waiting truck-available dock pairs (from the set of idle docks and set of trucks waiting in the system) satisfying $x_{t,d}$>0. Select the "waiting truck-available dock" pair having the minimum combined parameter. If there is a tie, select one of these candidates randomly.

  (iii) There is at least one truck waiting for unloading at dock *d* queue: Call the selected truck from the queue of the dock *d*.

- A Truck comes to the queue for unloading.

  (i) Dock to which the truck is assigned is busy and there is no other available dock for unloading or all possible “waiting truck-available dock” pairs have values of $x_{t,d}$=0: There is no action to take. The arriving truck enters to the queue and waits for unloading.

  (ii) Dock to which the truck is assigned is busy and at least one other dock satisfying $x_{t,d}$>0 is available for the unloading operation: Investigate all waiting truck-available dock pairs (from the set of idle docks and set of trucks waiting in the system) satisfying $x_{t,d}$>0. Select the “waiting truck-available dock” pair having the minimum combined parameter. Note that the only truck satisfying this condition can be the last arriving truck.

  (iii) Dock to which the truck is assigned is available: Send the arriving truck to the determined dock.

4.2.2.4. Static Allocation Policy with Attribute Dependent Admission under Threshold Queue Length Constraint (StAllc_THAdms). Another hybrid policy can also be constructed by unifying minimum combined parameter policy and static allocation policy. At the first stage, arriving trucks are assigned to docks according to static allocation policy. In this system, each unloading dock queue has its own threshold value. Due to the easiness of implementation and stochastic setting of unloading times, number of waiting trucks is used as a control variable for threshold policies. At a decision point, if at least one of the queues exceeds its threshold value, then rule selects the “waiting truck-available dock” pair having minimum combined parameter from the eligible queues (exceeding threshold level). Actually, when threshold levels are zero, this policy behaves as minimum combined parameter policy and when the threshold levels go to infinity, it converges to static allocation policy.

Using threshold policies in queuing system is a widespread approach in literature. As Vericourt and Zhou [5] and Mehrotra [11] state in their studies, threshold policies help to balance between different objectives. In our case, although static assignments of trucks are made at the arrival time of trucks, when queue lengths go up, admission decisions to other docks can be made to reduce waiting time via threshold policy.

Threshold values can be determined by decision maker according to the weighted importance of performance measures (waiting time and penalty rate). In this thesis, rules with threshold queue lengths 1, 2, 5 and 10 are simulated and resulting tradeoffs curves are analyzed for these different threshold values. The policy algorithm can be summarized as follows:

- An unloading operation is completed at dock *d*

    (i) There is no truck waiting for unloading and all unloading queue lengths are lower than their threshold values: There is no action to take. Available docks wait for truck arrival.

    (ii) There is no truck waiting for unloading and at least one unloading queue length exceeds threshold value: Investigate all waiting truck-available dock pairs from the set of idle docks and set of trucks waiting in the queues that exceeds the threshold value. Select the "waiting truck-available dock" pair having the minimum combined parameter. If there is a tie, select one of these candidates randomly.

    (iii) There is at least one truck waiting for unloading at dock *d* queue: Call the selected truck from the queue of the dock *d*.

- A Truck comes to the queue for unloading

    (i) Dock, to which the truck is assigned is busy and there is no other available dock for unloading or all unloading queue lengths are lower than their threshold values: There is no action to take. The arriving truck enters to the queue and waits for unloading.

(ii) Dock, to which the truck is assigned is busy and length of queue of dock *d* exceeds threshold value with the arriving truck and at least one of the other docks is available: Investigate all waiting truck-available dock pairs from the set of idle docks and set of trucks waiting in the queues that exceeds the threshold value. Select the "waiting truck-available dock" pair having the minimum combined parameter. If there is a tie, select one of these candidates randomly.

(iii) Dock, to which the truck is assigned is available: Send the arriving truck to the determined dock.

Considering the characteristics of the policies, it is seen that minimum unloading time policy is a waiting time centric policy while minimum penalty policy is a penalty rate centric one. Other policies can be stated as hybrid policies that are trying to balance different performance measure. In this study, policies are referred by using the abbreviations or policy numbers given in Table 4.1

Table 4.1. Simulated policies

Policy Number	**Abbreviation**	**Policy Name**
1	Rand	Randomized Routing Policy
2	Min_UT	Minimum Unloading Time Policy
3	Min_P	Minimum Penalty Policy
4	Min_CbPr	Minimum Combined Parameter Policy
5	StAllc	Static Allocation Policy
6	StAllc_ARAdms	Static Allocation Policy with Attribute Dependent Admission under Assignment Ratio Constraint
7 (TH)	StAllc_THAdms (TH)	Static Allocation Policy with Attribute Dependent Admission under Threshold Queue Length Constraint [(TH) : Threshold Value]

4.3. Simulation Models

ARENA Simulation program is used to develop simulation models for the explained routing policies. Seven different simulation models are constructed. Considering the model settings, there are two possible decision epochs in the system. The first epoch is the time when an arrival process occurs and the second epoch is the time when an unloading service process is completed. The algorithms of the simulation models are explained deeply in the following sections.

4.3.1. Single Queue Policies

In these kinds of policies arriving trucks are waiting in a single queue. At the decision epoch, one of the waiting trucks and one of the idle docks are selected to perform the unloading operation. Truck-dock selection can be made according to the three different single queue policies:

- Minimum unloading time policy
- Minimum penalty policy
- Minimum combined parameter policy

These simulation models mainly consist of two submodules. The first one represents the main processes where arrival, unloading and departure activities are represented. When a decision epoch occurs, a signal is send to second sub module where possible waiting truck-idle dock pairs are investigated.

After this investigation, one of the waiting trucks and one of the idle docks are determined according to the implemented policy. Afterwards determined truck is sent to determined dock and process goes on. The flowchart of the simulation model is given in Figure 4.1.

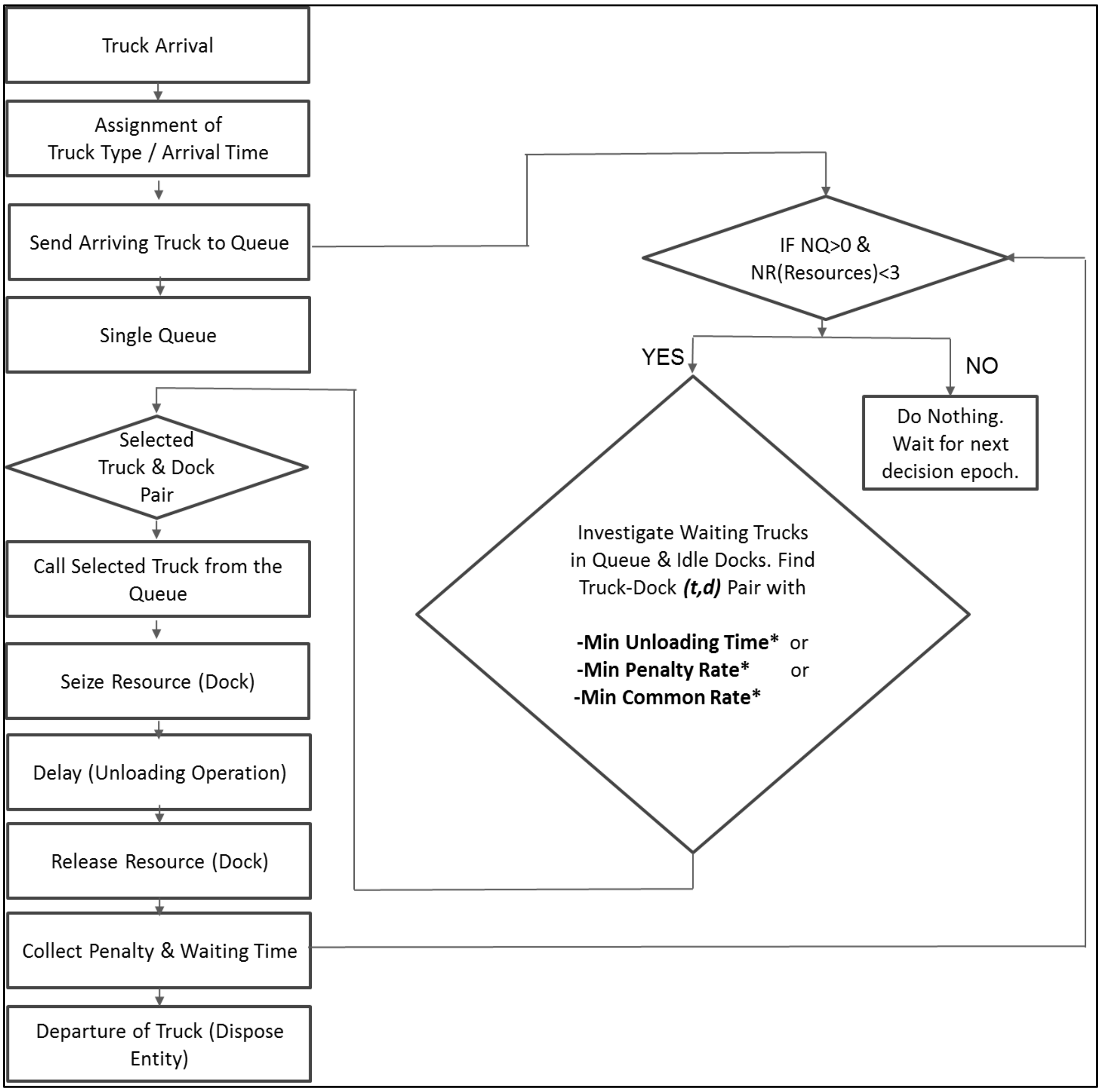

Figure 4.1. Simulation modules for policy 2, policy 3 and policy 4

4.3.2. Dock Dedicated Queue Policies

In these kinds of policies, arriving trucks are waiting in different dock dedicated queues. At the arrival instant of a truck is the first decision epoch where the dock-dedicated queue to which the arriving truck is directed is determined according to pre-determined assignment ratios. For randomized routing policy, these assignment probabilities are same while for the other dock-dedicated policies the assignment probabilities are determined by an optimization model explained in Chapter 4.2.2.

Four different dock dedicated queue policy is simulated:

- Randomized routing policy
- Static allocation policy
- Static allocation policy with attribute dependent admission under assignment rate constraint
- Static allocation policy with attribute dependent admission under threshold queue length constraint

For randomized routing policy and static allocation policy, trucks are waiting in the dock dedicated FCFS queues until the end of the unloading process and there is no any other decision epoch. However, other dock dedicated queue policies that allow to dynamic admission from other queues also has a decision epoch at the time when an unloading service is completed.

Simulation models constructed for the policies that allow admission from other dock-dedicated queues mainly consist of three submodules. The first one represents the main processes where arrival, unloading and departure activities are represented and the second one makes decision whether an admission from different queue should be made or not at the decision epochs. Lastly, the third sub-module is activated if the second one decides for admission from a different queue. Then, this sub-module determines which truck should be accepted and which dock is used for unloading operation. Since randomized routing rule and static assignment rule do not allow admission from different dock queue, only the first sub-module is used in the simulation models of these two policies. The flowchart representation for dock-dedicated policies are given in Figure 4.2 and Figure 4.3.

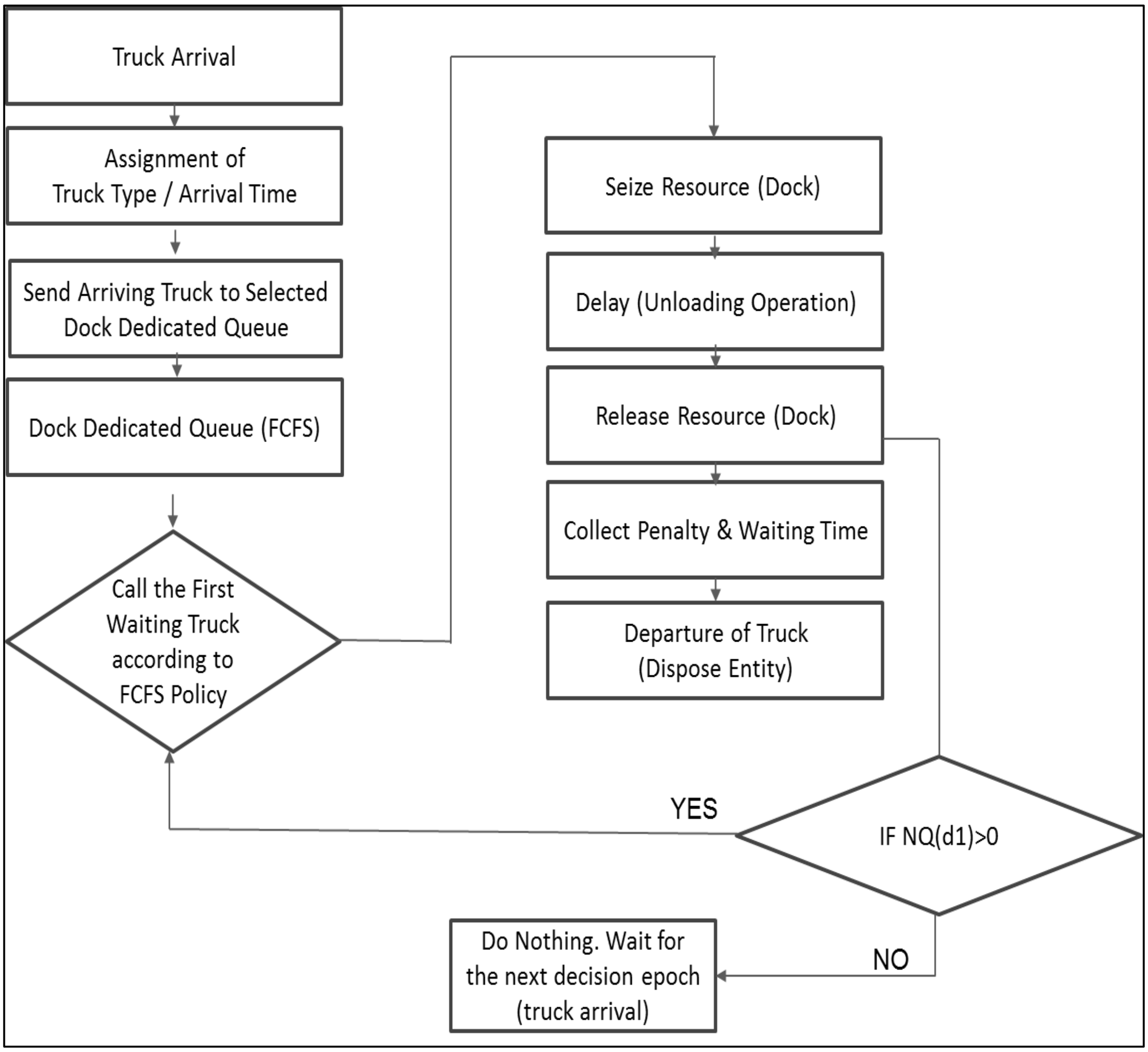

Figure 4.2. Simulation Modules for policy 1 and policy 5

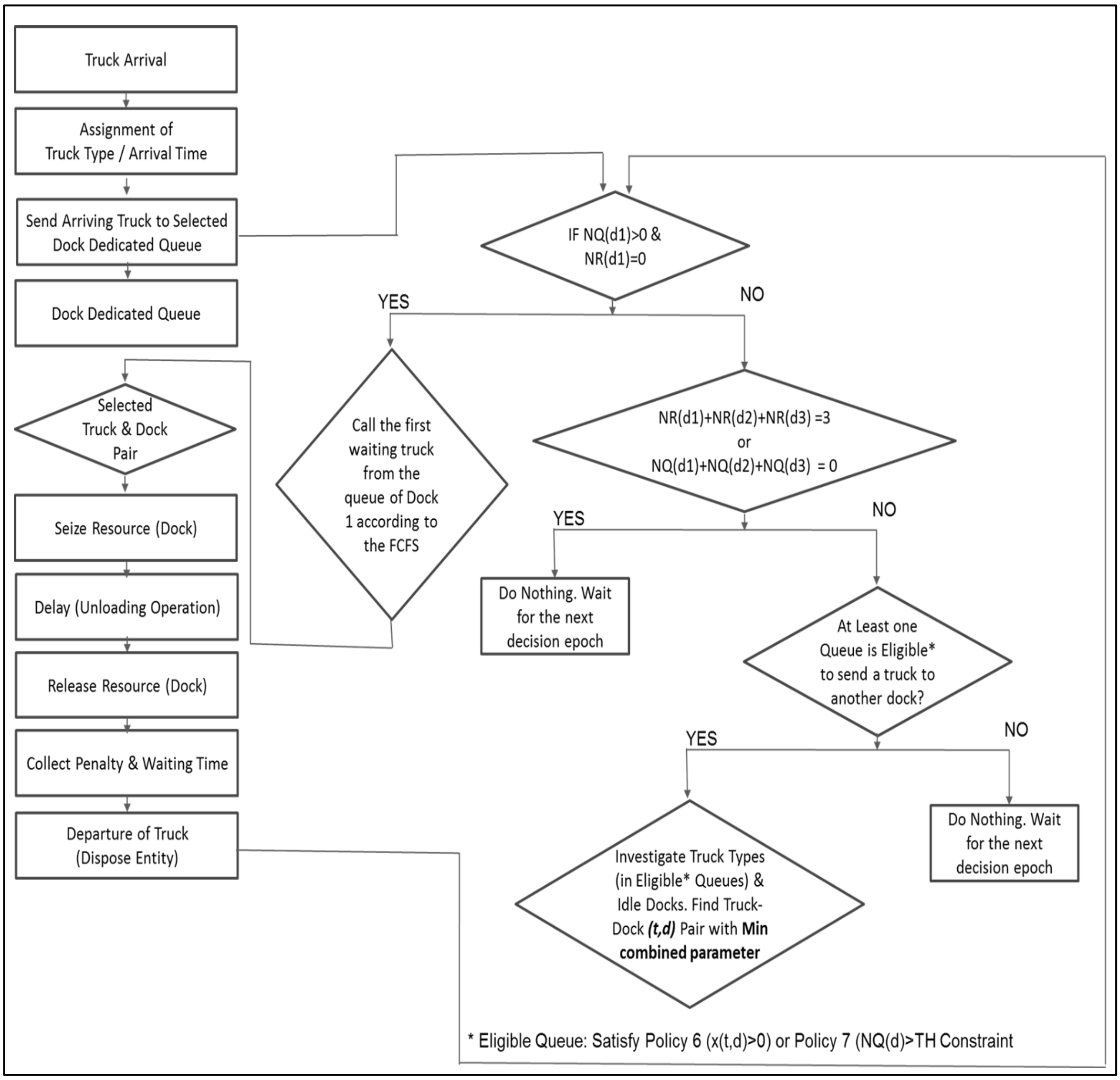

Figure 4.3. Simulation Modules for policy 6 and policy 7

5. DESIGN OF SIMULATION EXPERIMENTS AND SCENARIO ANALYSIS

After introducing different system settings and several routing policies, it is need to investigate the strengths and weaknesses of these policies. Since the problem definition includes a complex environment comprising multiple truck types and heterogeneous unloading docks, design of simulation experiments is a convenient method to observe the response of different policies in several environments. Without doubt, some policies are more effective when the system parameters are in between some specific values. To illustrate, certain policies perform well in the systems where traffic intensity is quite high while some others are more effective in light traffic environments.

Average waiting time per truck and average penalty rate are considered as two different response variables of the system. Therefore, there is no a certain answer about the selection of the best policy. Instead, at the end of the study, it is aimed to obtain some level of information about the effectiveness of simulated policies in different environments.

5.1. Design of Simulation Experiments

In order to simulate routing policies explained in Chaper 4, ARENA Simulation Software is used. All the representations used in model construction phase are given in the following sub-sections.

5.1.1. Resources

Unloading docks are the resources serving to arriving trucks. Three unloading docks are created in all simulation models. Representation of these docks in simulation models are as follows:

- Resource1 : Unloading dock 1
- Resource2 : Unloading dock 2
- Resource3 : Unloading dock 3

5.1.2. Queues

Single queue models have only one queue for all types of trucks while dock-dedicated queue models include one dock for each unloading docks. Following dock definitions are used in constructed simulation models:

- Dock1Q : Separated queue for dock 1
- Dock2Q : Separated queue for dock 2
- Dock3Q : Separated queue for dock 3
- QueueQ : Shared queue for single queue policies

5.1.3. Response Variables

"Average Waiting Time of Trucks" and "Average Penalty Rate" are the response variables of the system. Performance of different routing policies are compared with each other by investigating these values.

5.1.4. Factors and Levels

The output of a specific policy simulation can be affected by different levels of system parameters. Each policy is simulated for different levels of the factors.

<u>5.1.4.1. Traffic Intensity Rate.</u> Arrival and service rates determine the traffic intensity level of the system. It mainly represents the average utilization levels of unloading docks. Traffic Intensity is simulated in three different utilization level. Low (65%), Normal (75%) and High (85%) utilization level scenarios are constructed for the experimental study. Expected traffic intensities are measured based on randomized routing rule scenarios for each level.

$$Utilization_d = \frac{\Sigma_t(\lambda_t\,/\mu_{t,d})}{3} \tag{5.1}$$

5.1.4.2. Correlation between service time and penalty rate parameters. In our system settings, each truck type-dock pair has an expected service time and penalty rate. It is expected that the correlation between expected service time "$1/\mu_{t,d}$" and penalty rate "$p_{t,d}$" values have a significant effect on the performance of different routing rules.

Positive correlation implies that dock *d* has low (high) "$1/\mu_{t,d}$" and low (high) "$p_{t,d}$" for the type *t* of trucks. Therefore, with respect to both service time and penalty rate dock *d* is desirable (undesirable) for truck type *t*.

On the other hand, if the dock *d* has high (low) "$1/\mu_{t,d}$" and low (high) "$p_{t,d}$" for type *t* of trucks, they are called as negatively correlated. With respect to either service time or penalty rate aspect, dock *d* is preferable but for the other one it is not.

Considering the correlation scenarios, three input levels are used: Positive Correlation (+0.8), Negative Correlation (-0.8) and No Correlation (0) between service rates and penalty values. Correlation coefficient between unloading times and penalty rates are calculated by:

$$Correl\ (\mu, p) = \frac{\sum_{t,d}(1/\mu_{t,d} - 1/\mu_{avr})(p_{t,d} - p_{avr})}{\sqrt{\sum_{t,d}\left(1/\mu_{t,d} - 1/\mu_{avr}\right)^2 \sum_{t,d}\left(p_{t,d} - p_{avr}\right)^2}} \tag{5.2}$$

5.1.4.3. Dispersion of penalty rates. For a truck type *t*, each dock has a penalty rate $p_{t,d}$. Gaps between these penalty values for a specific truck type affect the effectiveness of the routing policies in different scenarios. For a specific truck type *t*, penalty levels of different docks can be very similar (even same) or can be far from each other. For this factor, coefficient of variation is used as a measure of dispersion. Two levels of parameter sets are simulated for dispersed penalty rates (CV: 0.7) and gathered penalty rates (CV: 0.1). Coefficient of variation measures for penalty rates are calculated as follows:

$$CV_{system} = \frac{\sum_t \lambda_t \, CV_t}{\sum_t \lambda_t} \tag{5.3}$$

$$CV_t = \frac{\sigma_{(t,d)}}{\mu_{(t,d)}} \quad for\ t = 1{,}2{,}3 \tag{5.4}$$

5.1.4.4. Scenarios. Considering the levels (3x3x2) of abovementioned factors, 18 different scenarios are constructed to test seven different routing policies. Each routing policy is simulated for 18 different scenarios. These scenarios are differentiated according to the levels of three different factors that are seen in Table 5.1. All different factor levels are given in Table 5.2, Table 5.3 and Table 5.4.

- Traffic Intensity Rate (High, Normal, Low)
- Correlation between service time and penalty parameters (Negative, Zero, Positive)
- Dispersion of Penalty Rates (High, Low)

Table 5.1. Levels of three different factors

Scenario No	**Traffic Intensity**	**Correlation Between Expected Service Time and Penalty Rate**	**Dispersion of Penalty Rates**
1	Low	No Correlation	Low
2	Low	Negative Correlation	Low
3	Low	Positive Correlation	Low
4	Low	No Correlation	High
5	Low	Negative Correlation	High
6	Low	Positive Correlation	High
7	Normal	No Correlation	Low
8	Normal	Negative Correlation	Low
9	Normal	Positive Correlation	Low
10	Normal	No Correlation	High
11	Normal	Negative Correlation	High
12	Normal	Positive Correlation	High
13	High	No Correlation	Low
14	High	Negative Correlation	Low
15	High	Positive Correlation	Low
16	High	No Correlation	High
17	High	Negative Correlation	High
18	High	Positive Correlation	High

Table 5.2. Traffic intensity levels

Traffic Intensity Levels	**Values**
High	85%
Medium	75%
Low	65%

Table 5.3. Correlation levels

Correlation Between Expected Service Time and Penalty Rate	**Values**
No Correlation	-0.8
Negative Correlation	0
Positive Correlation	+0.8

Table 5.4. Dispersion of penalty levels

Dispersion of Penalty Rates	**Values**
High	0.7
Low	0.1

5.2. Output Analysis

The most important performance measures of the system are average waiting time and average penalty rate per a truck. As explained previous chapters, penalty rate represents the distance covered to carry a truckload from unloading dock to the warehouse. Since the relative importance of these two performance measures can change in different organizations/system settings, it is very difficult to quantify the weights of these outputs by creating weight coefficients. Instead, in this thesis, both performance measures are obtained separately and integrated output analysis are performed by using trade-off curves. Depending on the input parameters (penalty rates and unloading times), two performance measures may be conflicting in some system settings. For instance, for the environment where unloading times and penalty rates are negatively correlated, faster docks can be far from the warehouse of the materials. Therefore, selecting faster docks reduces unloading times and waiting times of the trucks while the distance covered to carry a truckload from unloading dock to the warehouse (penalty rate) is excessive due to this selection.

All performance measures are analyzed at steady state of the system. Detailed steady state analyses are explained for the first model (Scenario 1) that simulates randomized policy. Output analysis of all other models are performed in the same way. All the analysis as warm-up period or replication number determinations are studied with the methodology explained in the book of Low and Kelton [23].

5.2.1. Warm-Up Period and Replication Length

Warm-Up period represents the time past until the system reaches steady state. In order to get reliable results, this time interval is cut and the output analyses are made by observing the data collected in steady state phase. Additionally, replication lengths of models are determined as 4 times warm-up period intervals. Warm-Up period/replication length determination processes are implemented for all simulation scenarios. As an example, the steps of the study for Scenario 1 can be summarized as follows:

- Using the Output Analyzer module of ARENA software, cumulative average plot is drawn for occurred waiting time in the simulation model. Cumulative average plot shows the waiting time average of all trucks that are unloaded until that time. When the cumulative average curve remains stable as new realizations are occurring, it can be asserted that the system is working in steady state.

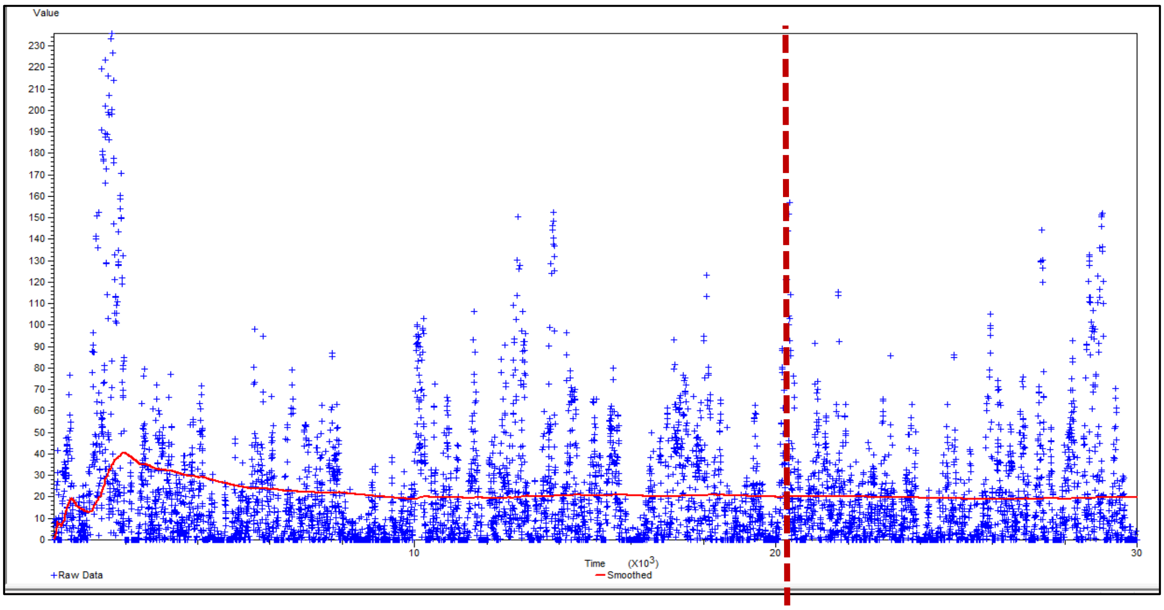

Figure 5.1. Determination of warm-up period for randomized policy with low utilization, zero correlation and low dispersion scenario. (scenario 1, replication 1)

- Investigating cumulative average plot (example is given in Figure 5.1) , it is observed that average waiting time per truck reaches steady state after 20.000 minutes. Thereby, first 20.000 minutes should be truncated as Warm-Up period. After that, we preferred a replication length that is 4 times longer than truncated part to gain more reliable outputs [23]. Therefore, replication length is determined as 80,000 minutes and warm-up period as 20,000 minutes. By this way, we can get reliable data for 60,000 minutes that is three times of truncated point.

5.2.2. Replication Number

Similar to warm-up and replication length determination, required replication number of a simulation scenario is also determined. The proses mainly aims to decrease half-width for the 95% confidence interval as small as the 20% of the average mean value. The proposed procedure for the replication number determination is given in Figure 5.2.

S0. Make $n = n_0$ initial replications. (n_0 is usually a small number like 5.)

S1. Find $\bar{\bar{Y}}$, s and h using n replications made. Choose desired half width h^*. Given the desired relative precision, r, (say r = 0.20), desired half-length can be determined as $h^* = r \times \bar{\bar{Y}}$.

S2. If $h^* \leq h$ then stop with CI $\bar{\bar{Y}} \pm h$.

Otherwise, find $n^* = [n(h/h^*)^2]+1$. Alternatively use $n^* = n + 1$.

Make $n^* - n$ additional replications (or make n^* replications from scratch).

Set $n^* = n$.

Go back to S1.

Figure 5.2. Procedure for determination of replication number given by [23]

To illustrate, for randomized policy with low utilization, zero correlation and low dispersion scenario (Scenario 1), the confidence intervals are constructed as shown in Table 5.3

Table 5.5. Half-width/average ratios and replications for randomized rule

	AVERAGE	0.950 C.I HALF-WIDTH	RATIO	NO OF OBS.
TAVG(WaitingTime)	29.8	2.16	0.0725 -OK	5
TAVG (RealPenalty)	62.7	0.0735	0.0012 -OK	5
DAVG (NumberInQ.)	4.3	0.325	0.0756 -OK	5
DAVG(Resource1.Utilization)	0.63	0.0106	0,0173 -OK	5
DAVG(Resource2.Utilization)	0.599	0.00906	0,0151 -OK	5
DAVG(Resource3.Utilization)	0.752	0.0132	0,0175 -OK	5

“Half-width/Average” ratios are calculated and the 95% precision levels are found for each output factor. As shown in table above, relative precision levels are lower than 0.2 in five replications. Thus, five replications are sufficient to gain reliable results. If these ratios were bigger than 0.2 considerably we would apply the iterative process to find appropriate replication number *n* by using $n^* = [n(h/h^*)2]+1$ where h is half-width as stated in Figure 5.2. The procedure is implemented in all simulation scenarios. The required number of replication, warm-up periods and replication lengths of each run is determined.

5.2.3. Verification and Validation

After the construction of simulation models, verification and validation studies are performed. Verification phase checks whether the conceptual model is translated into a software or a digital media correctly or not. On the other hand, validation step is related to see if the simulation model represents the real life system to reach compatible process results.

<u>5.2.3.1. Verification of Models.</u> Designing correct simulation models are one of the most vital steps of our study. Due to the complexity of the models with many inputs, variables and submodules, it is quite hard to detect the interactions of these components and to be sure about verification of the models.

To verify constructed simulation models, several controls are performed during the verification process.

- After the logical investigation of the models, TRACE element is used in ARENA. By this way, model operations are traced and the correct entity flows during the operations are verified. (Ex: for the arrival process, it is traced that first entity come into the line and then searching operation is started)

- Different input rates are tried and outputs are investigated for different input parameters. (Ex: it is seen that waiting times and utilization rates of servers are much higher than when arrivals rates are increased or service rates are decreased.)

- The output statistics are compared with each other and consistency between results is assured. (Ex: in simulation runs in which waiting time statistics are high, number in queues values and average utilization rates of servers are also high)

- Simulation sub-models are investigated separately and it is ensured that each sub-model runs correctly. (Ex: main arrival-service process sub-module and truck-dock selection sub-modules are traced independently and the separate results of the sub-models are observed)

- Counters and control variables are used to check the balance of the entities at the end of the simulation runs. (Ex: Created(JIS) + Created(Domestic) + Created(Import) = Serviced(InResource1) + Serviced(InResource2)+ Serviced(InResource3) +NQ(Q1) + NQ(Q2) + NQ(Q3) + NR(Server1) + NR(Server2) + NR(Server3))

5.3.2.2. Validation of Models. Since the constructed models simulate experimental settings with different input parameters, it is not possible to compare them with real life reference cases. In such studies, comparing theoretical results and simulation outputs is a powerful method to validate experimental models. For homogenous customers and homogenous servers environment with exponential inter-arrival and service times, analytical formulation for average waiting times per customer is known. To assure the validity of constructed simulation models, all truck types and dock types are similarized with same arrival and service time parameters. All trucks are assumed to be single type and inter-arrival times are same as the inter-arrival time of truck type 1 (Expo (15) like in Scenario 1). Service time for each truck is also commonized as Expo(9) for all docks like in Scenario 1. In other words, simulation models are run with one type of truck and 3 same unloading docks as M/M/c (where c=3) model setting. Than expected waiting time per truck is calculated by using little's formula ($l = \lambda w$) with queue length formulation for M/M/c model as:

$$L_q = \frac{P_0 \left(\frac{\lambda}{\mu}\right) \rho}{c!(1-\rho)^2} \tag{5.5}$$

where

$$P_0 = \sum_{m=0}^{c=1} \frac{(c\rho)^m}{m!} + \frac{(c\rho)^c}{c!(1-\rho)} \quad (5.6)$$

$$\rho = \frac{\lambda}{c\mu} \quad (5.7)$$

P_0 donates the probability that there is no truck (customer) in the system and c donates the number of docks (servers).

$$\text{Wait in the Queue} = W_q = \frac{Lq}{\lambda} \quad (5.8)$$

Then it is seen that, calculated analytical result is between high and low limits of 95% confidence interval obtained from simulation outputs as shown in Table 5.6

Table 5.6. Comparison of analytical results and simulation outputs for waiting time of trucks for M/M/3 queue

	Analytical Result W(q) (min)	**Simulation Output 95% CI W(q) (min)**
Waiting in Queue	2.66	[2.54 , 2.81]

As explained in model development chapter, static allocation policy includes different dock dedicated queues. Arriving trucks are assigned to dock dedicated queues according to the pre-determined assignment ratios ($x_{t,d}$) and FCFS discipline is implemented for each queue. Therefore, we can investigate the system as three different M/M/1 queuing systems when all *t*-type trucks are assigned the dock *d* (where only one $x_{t,d} = 1$ for all t. Hence, one-to-one truck type-dock matching for truck allocation is achieved*)*. Using the expected inter-arrival times and expected service time of each dock; expected waiting time per truck can be calculated by using analytic formulae for each M/M/1

To illustrate consider scenario 6 where traffic intensity is low with positive correlation and high level of dispersion. Resulting $x_{t,d}$ values used in static admission rule is as follows:

Table 5.7. Assignment ratios ($x_{t,d}$ values) obtained from optimization model of static admission rule in scenario 6

$x_{t,d}$		Dock (d) 1	Dock (d) 2	Dock (d) 3
Truck (t)	1	**1**	0	0
	2	0	**1**	0
	3	0	0	**1**

Therefore, three different M/M/1 systems appear for each dock where all type 1 trucks enter the queue of dock 1, all type 2 trucks enter the queue of dock 2 and all type 3 trucks enter the queue of dock 3.

Comparison of analytic results and simulation outputs shows that analytical results drop inside 95% confidence interval for waiting in queue time that is obtained from the simulation study as seen in Table 5.8.

Table 5.8. Comparison of analytical results and simulation outputs for waiting time of trucks for three different M/M/1 queue.

	Analytical Result W(q) (min)	**Simulation Output 95% CI W(q) (min)**
Queue1	13.5	[12.3 , 15.1]
Queue2	13.7	[12.3 , 15.7]
Queue3	8	[7.88 , 8.54]

Another validation check for the previous model can be performed by using exponential inter-arrival time and lognormal service time distributions. In this case, each queue becomes an independent M/G/1 queue and Pollaczek-Khinchin (P-K) formula can be used to see expected waiting time for each dock dedicated queue. Comparing the analytical results and simulation outputs, it is concluded that the theoretical values are between the upper and lower bounds of 95% confidence intervals constructed by using simulation outputs as seen in Table 5.9.

Pollaczek-Khinchin (P-K) Formula:

$$W = \frac{\lambda E[X^2]}{2(1-\rho)} \tag{5.9}$$

where

$$\rho = \lambda/\mu = \lambda E[X] \tag{5.10}$$

$$E[X^2] = \mu^2 + \sigma^2 \tag{5.11}$$

Table 5.9. Comparison of analytical results and simulation outputs for waiting time of trucks for three different M/G/1 queue.

	Analytical Result W(q) (min)	**Simulation Output 95% CI W(q) (min)**
Queue1	10.125	[9.25 , 11.5]
Queue2	10.286	[9.04 , 10.5]
Queue3	6	[5.46 , 6.48]

5.3. Numerical Results

By simulating 18 different scenarios, performances of seven different routing policies in different environments are observed. Three environmental factors that have effects on performances of routing policies are as follows:

- Traffic intensity rate
- Correlation between service time and penalty parameters
- Dispersion of penalty rates

Those environmental factors are controlled during the simulation runs and results are evaluated by comparing outputs of several scenarios. By investigating the outcomes of simulation models, resulting performance measures are obtained for different environments. Figures 5.3-5.8 show the results of different scenarios with different factor levels and indicate routing policies included by efficient frontier borders. The policies that are not in the efficient frontier border are dominated by at least one other policy. Results of 18 different scenarios and efficient frontier of each one is shown in the following graphs.

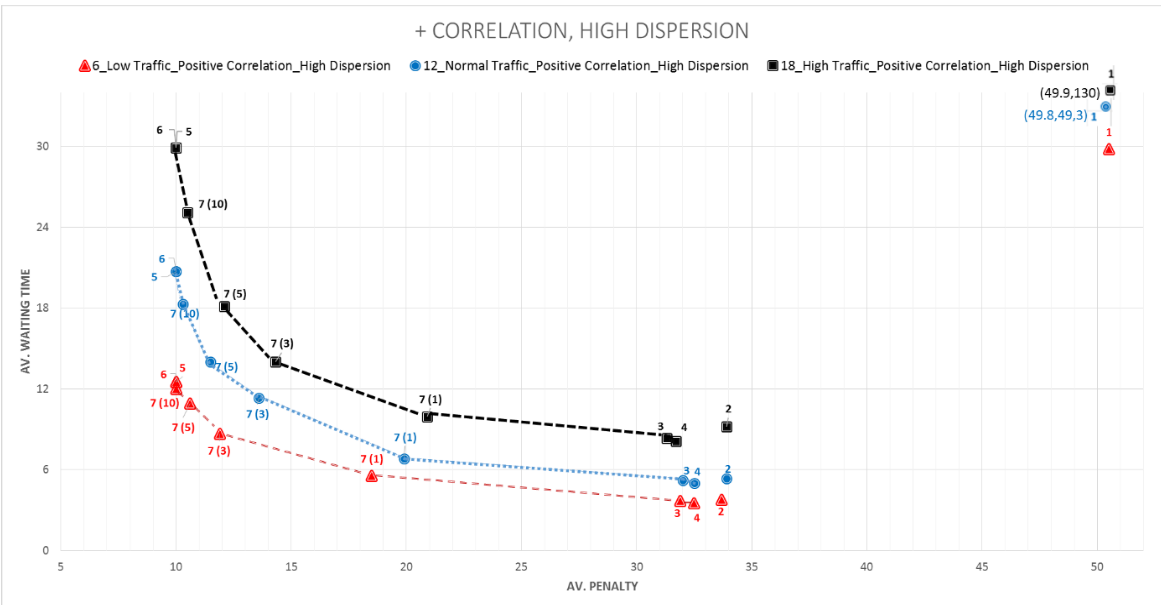

Figure 5.3. Outputs of high, normal, low traffic scenarios in positive correlation and high dispersion environment

Figure 5.3 demonstrates the results in environment where penalty rates and unloading times are positively correlated and penalty rates are highly dispersed. Policy 5 and policy 6 are the most efficient policies for penalty minimization aspect but average waiting times of these policies are quite high in all traffic intensities. On the other hand, policies 3 and 4 are located at the other end of the frontier where average waiting times are minimum but penalty rates are quite high.

It is seen that policies 2,3 and 4 give very near results due to the positive correlation between penalty and waiting times. Since penalty rates and unloading times are positively correlated, any of these policies aiming to minimize average penalty rate (average waiting time) is also efficient to minimize average waiting time (average penalty rate). For this environment, policy 1 is far from the efficient frontier line and dominated by other policies. Therefore, allocating coming trucks to queues randomly, results in very high waiting time and penalty rate.

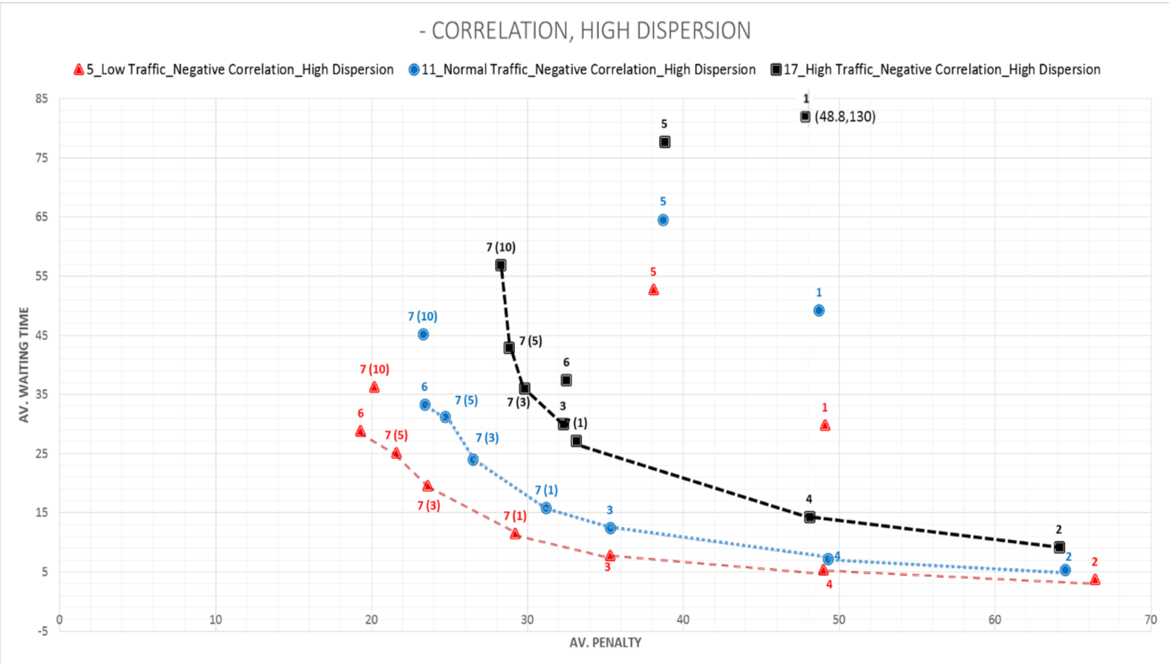

Figure 5.4. Outputs of high, normal, low traffic scenarios in negative correlation and high dispersion environment

Negative correlation and high dispersion environment results are given in Figure 5.4. Compared to positively correlated environment, both average waiting times and penalty rates are higher in this environment. Moreover, policy 5 moves to outside of the efficient frontier. The reason is related to the static characteristics of policy 5. As mentioned in Chapter 4.2, truck assignment ratios are determined by using 0.85 utilization level constraint for each dock. Due to this constraint, some portion of trucks cannot be assigned to their ideal docks in order not to violate 0.85 utilization level constraint. Hence, this situation results in some level of inefficiency. Additionally, during the operations, even if a dock is idle, policy 5 does not allow to serve a truck from another dock's queue. Therefore, policy 5 misses the opportunity to serve some trucks in their ideal docks even if it is idle at that time. This situation has much drastic effects in negative correlation and high traffic levels as seen in Figure 5.4. Due to the negative correlation between waiting times and penalty rates, two performance measures are conflicting in this environment. Therefore, the policies gives minimum waiting time results in excessive penalty rates or vice versa. Hence, waiting time-penalty balancing rules like policy 4 or policy 7(1) can find an acceptable balance point.

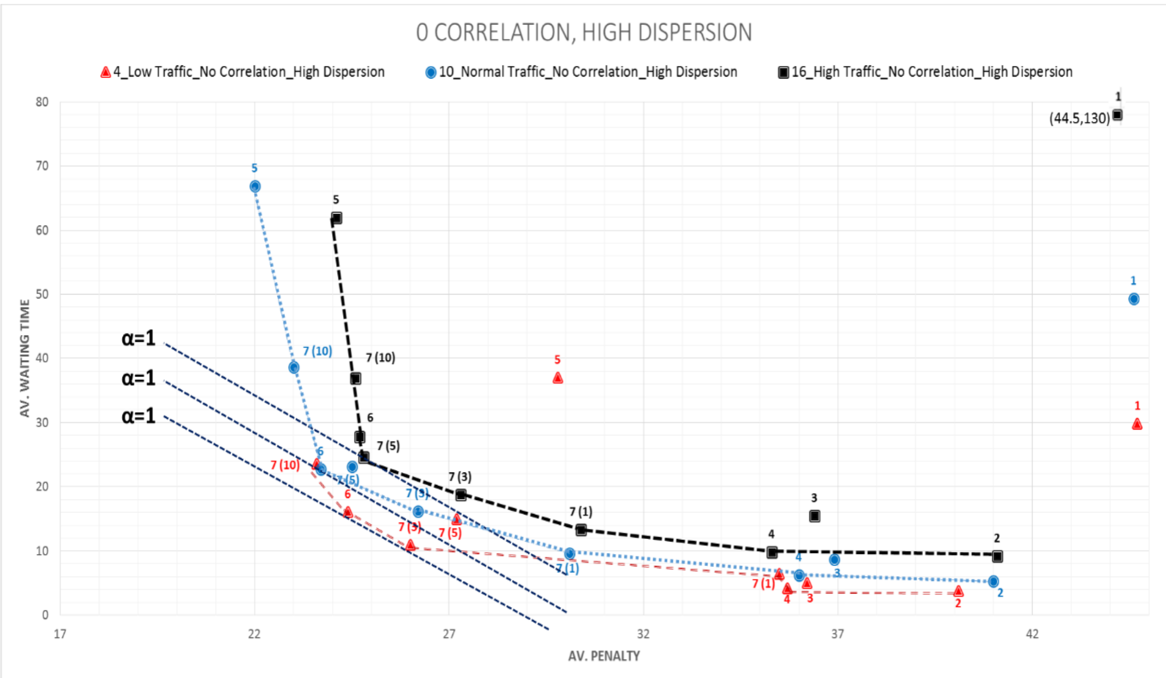

Figure 5.5. Outputs of high, normal, low traffic scenarios in zero correlation and high dispersion environment

Figure 5.5 shows the outputs of the policies in the environment where unloading times and penalty rates are uncorrelated. Actually, this is the situation between the previous two environments. Although output points are closer to each other compared to negatively correlated environment, again using mix-balancing hybrid rules (like policies 4,6 or 7) seems reasonable to balance excessive waiting times and penalty rates.

In order to determine which policies are highly efficient to satisfy both penalty and waiting time minimization aims, trade-off curves can be used during the graphical analysis. According to relative importance of these two performance measures, α-curves (relative importance curves) can be constructed and policies that are located near to this curves are determined as efficient policies. For instance, if relative importance of 1 unit of waiting time compared to 1 unit of penalty is equal to α than $\frac{Average\ waiting\ time}{Average\ penalty\ rate} = \alpha$ curves can be drawn. To illustrate if we take $\alpha = 1$ in Figure 5.5, best policies that are very near to α=1 curves in different environments would be plotted as follows:

- High Traffic Environment : Policy 7(5)
- Normal Traffic Environment : Policy 6
- Low Traffic Environment : Policy 7(3)

Determination of α is quite difficult and it is a parameter decided by system organizer.

Figure 5.6, Figure 5.7 and Figure 5.8 show the conjugate environments of previous three graphs. Similar results are also observed for this environment where the penalty dispersion levels are low. As a different observation, it is seen that efficient frontier lines are steeper in low dispersion environments compared to high dispersion ones. It means that, average waiting time can be improved considerably by sacrificing a small level of penalty in low dispersion environment. Since penalty rates are very close to each other, using waiting time centric rules does not result in very excessive level of penalty in low dispersion environment.

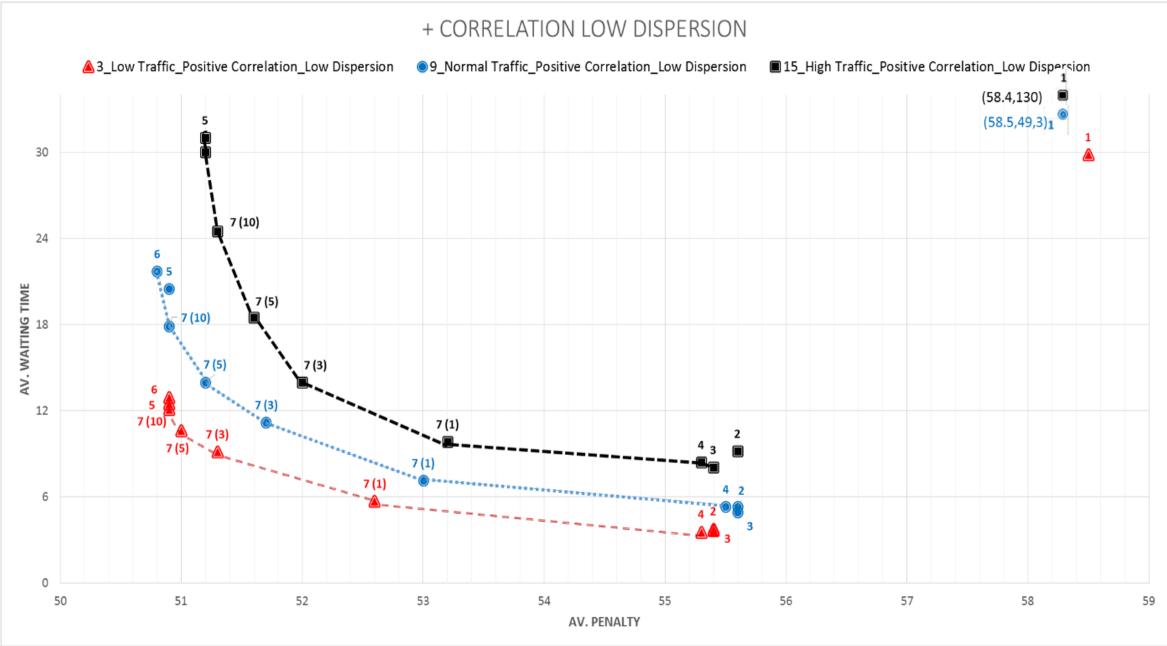

Figure 5.6. Outputs of high, normal, low traffic scenarios with positive correlation and low dispersion

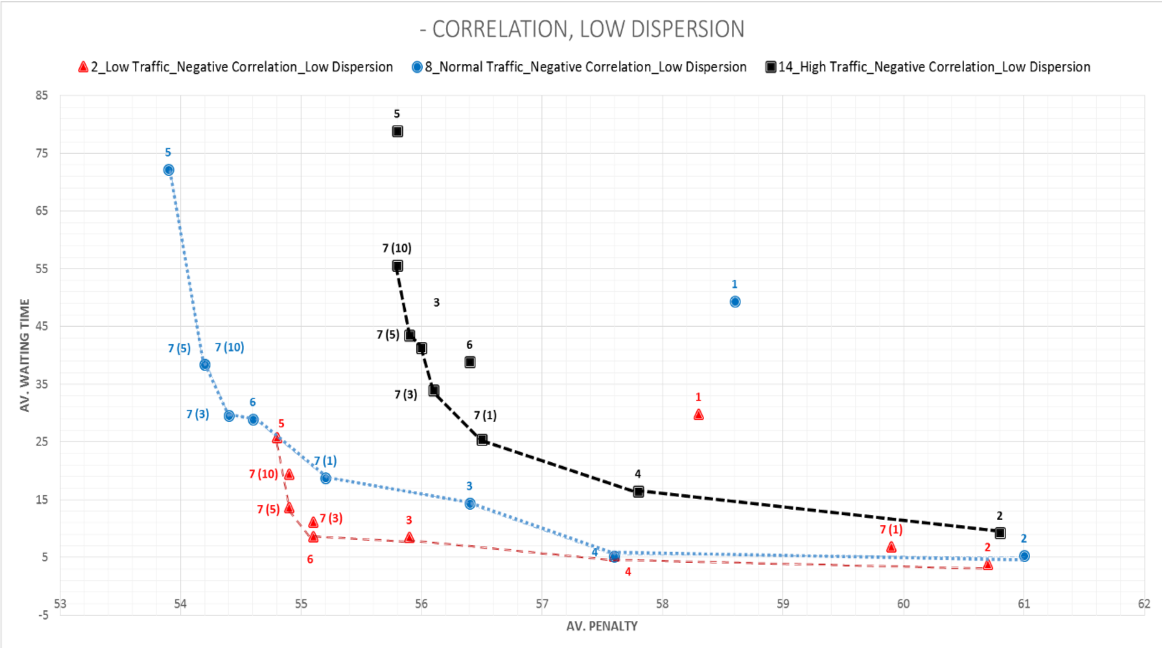

Figure 5.7. Outputs of high, normal, low traffic scenarios with negative correlation and low dispersion

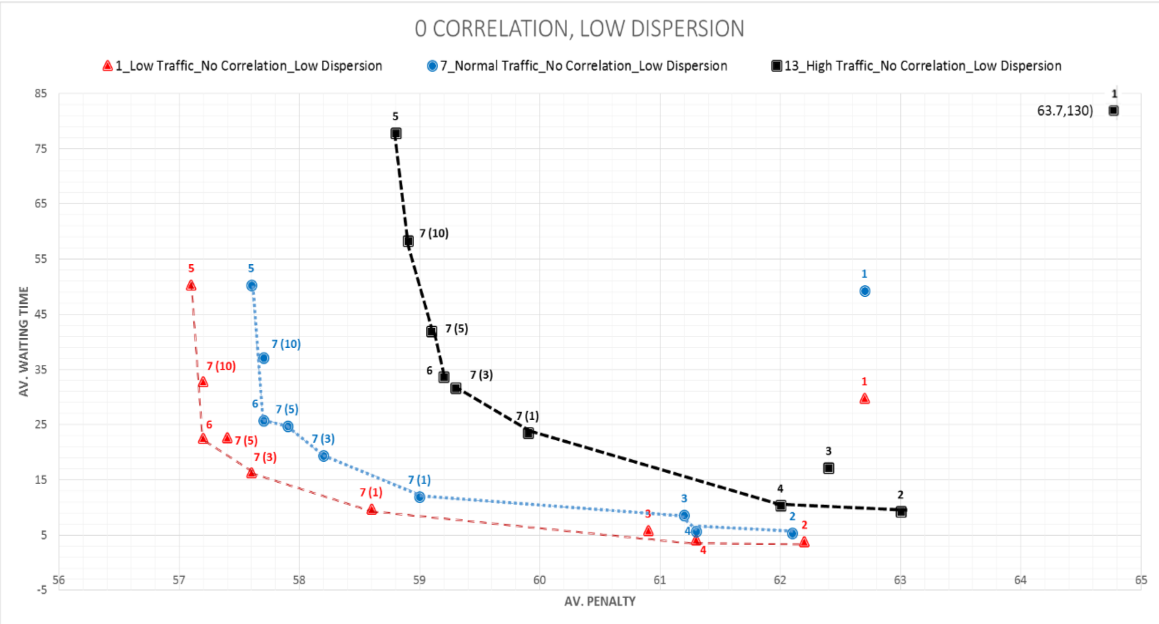

Figure 5.8. Outputs of high, normal, low traffic scenarios with zero correlation and low dispersion

5.3.1. Traffic Intensity Rate

Traffic intensity rate is one of the factors that influences the relative performances of different routing policies. As the traffic intensity goes up, both average waiting time and average penalty rate per truck increase. With the increase in traffic rate, the idleness of docks reduce and the probability of served by the ideal unloading dock for an arriving truck decreases. As a result, average queue lengths, waiting times and penalty rates are observed higher in the scenarios where traffic intensity rate is high. Investigating the graphs from Figures 5.3 to 5.8 it is observed that high traffic curves are always above the normal traffic curves for all scenarios. That indicates the higher level of average waiting time and penalty rate for high intense environments.

As explained in Chapter 4.2, static rules or rules with high level of thresholds [Rand, StAllc, StAllc_THAdms(5), StAllc_THAdms(10)] miss the opportunity to use idle unloading docks because trucks are assigned to dock-dedicated queues according to the static assignment rates. After that, assigned trucks are waiting for either the completion of their unloading operations (in StAllc policy) or satisfaction of threshold levels of the queues (in

StAllc_THAdms policies). Therefore, using static rule or rules with high threshold levels results in excessive waiting times. In high traffic environments, the problem is not very harmful because almost all docks are always busy during the day and there is already no opportunity to accept a truck from a different dock's queue. However, especially in the environments where the traffic intensity is low, many docks remain idle for long time periods but these docks cannot accept a truck from another dock dedicated queue due to the policy constraints. This situation makes these kinds of static policies relatively inefficient compared to other ones.

To illustrate, relative performance of StAllc_THAdms(10) policy against Min_UT, Min_P and Min_CbPr policies in high and low traffic intense systems can be investigated. Comparing the waiting time output ratios ["(StAllc_THAdms(10)/Min_X Policies)"] of scenarios 4 (low traffic) and 16 (high traffic), it is seen that the ratios in scenario 4 are higher whereas penalty ratios are quite similar. (it is seen in Table 5.10 and Table 5.11) Therefore, StAllc_THAdms(10) policy result in relatively worse waiting times performance compared to single queue policies (Min_X Policies). That is, the relative performance of static and high-level threshold policies compared to single queue policies or state based admission policies are worse in the environments where traffic intensity rate is low.

Table 5.10. Relative waiting time ratios StAllc_THAdms (10) policy to Min_X policies

Waiting Time Ratios (StAllc_THAdms(10)/Min_X)			
Scenarios	StAllc_THAdms(10) / Min_UT	StAllc_THAdms(10) / Min_PT	StAllc_THAdms(10) /Min _CbPr
4_Low Traffic_No Correlation_High Dispersion	6,29	4,69	5,67
16_High Traffic_No Correlation_High Dispersion	4,00	2,38	3,76

Table 5.11. Relative average penalty ratios of StAllc_THAdms (10) policy to Min_X policies

Penalty Ratios (StAllc_THAdms(10)/Min_X)			
Scenarios	StAllc_TH Adms(10)/ Min_UT	StAllc_TH Adms(10)/ Min_PT	StAllc_TH Adms(10)/ Min _CbPr
6_Low Traffic_Positive Correlation_Low Dispersion	0,59	0,59	0,59
21_High Traffic_Positive Correlation_Low Dispersion	0,60	0,68	0,70

StAllc_THAdms policies with low threshold levels give better penalty rates and quite good waiting time results where the traffic intensity is low. On the other hand, when traffic intensity of the system increases, these rules cannot perform as intended. The main reason is related to the queue lengths. When traffic intensity of the system is high, queues reach the threshold values in a short time and after this point; coming trucks are unloaded at the other unloading docks that are not the first choice for them. Since queues reach their threshold values in a very short time, using threshold rules would be meaningless. In other words, optimal assignment ratios deviate very much after a very short time (after the threshold value is reached) and high level of penalty values occurs when the threshold levels are low in highly intense environments. Therefore, low-level threshold policies should be preferred in the environments where traffic intensity rate is not too high. To illustrate, the range of resulting waiting time values are wider in low-intense traffic environments compared to high intense ones. It suggests that decision makers can get the desired output values by arranging threshold values in low-traffic environments easily. Considering the Figure 5.6, for instance, penalty values of policies 7(1) or 7(3) in high traffic environment are closer to penalty values of policy 4 compared to low traffic environment. Therefore, as traffic intensity increases, StAllc_THAdms policies with low thresholds values reach the threshold limits quickly and they approach to Min_CbPr policy.

Another important observation about StAllc_THAdms policies is that they are highly effective for low-intense environments with lower threshold values considering the balance of two different performance measures. Actually, StAllc_THAdms policies are derived from StAllc policy by relaxing dock dedicated unloading operation constraint. These policies allow docks to serve trucks waiting in different docks' queues if the lengths of these queues exceed the threshold level. This relaxation gets the opportunity to decrease excessive waiting times of trucks in exchange for some level of increase in penalty rates. It is seen that average penalty rate per truck is minimum when StAllc policy is implemented. By using lower threshold values, average penalty rates improve significantly at the cost of a relatively small increase of waiting time. As the threshold value increases, the cost of decreasing average penalty becomes highly expensive in terms of average waiting time and this effect gets larger as traffic intensity increases. To illustrate, scenario 13 (High Traffic) and scenario 1 (Low Traffic) can be compared.

Investigating Table 5.12 and Table 5.13, the first observation is that penalty rates decrease as threshold values increase. However, this improvement rate decreases as threshold value grows. Hence, increasing threshold values after a level cannot provide significant amount of penalty decrease. It only increase waiting times without sufficient improve in penalty rates. The second observation is that the mentioned effect is more explicit in high traffic environments. Therefore, it can be asserted that threshold policies are more efficient in low traffic intensity cases and with low threshold values.

Table 5.12. Traffic intensity effect on waiting time and average penalty tradeoff in high intense environment

<table>
<tr><th colspan="6">13_High Traffic_No Correlation_Low Dispersion</th></tr>
<tr><th>Policy</th><th>Av. Waiting Time</th><th>Av Penalty Rate</th><th colspan="3">Cost of decreasing average penalty 1 unit (in terms of waiting time)</th></tr>
<tr><td>StAllc_THAdms (TH : 1)</td><td>23,5</td><td>59,9</td><td rowspan="2">-13,67</td><td></td><td></td></tr>
<tr><td>StAllc_THAdms (TH : 3)</td><td>31,7</td><td>59,3</td><td rowspan="2">-51,00</td><td></td></tr>
<tr><td>StAllc_THAdms (TH : 5)</td><td>41,9</td><td>59,1</td><td></td><td rowspan="2">-82,00</td></tr>
<tr><td>StAllc_THAdms (TH : 10)</td><td>58,3</td><td>58,9</td><td></td><td></td></tr>
</table>

Table 5.13. Traffic intensity effect on waiting time and average penalty tradeoff in low intense environment

<table>
<tr><th colspan="6">1_Low Traffic_No Correlation_Low Dispersion</th></tr>
<tr><th>Policy</th><th>Av. Waiting Time</th><th>Av Penalty Rate</th><th colspan="3">Cost of decreasing average penalty 1 unit (in terms of waiting time)</th></tr>
<tr><td>StAllc_THAdms (TH : 1)</td><td>9,7</td><td>58,6</td><td rowspan="2">-6,60</td><td colspan="2"></td></tr>
<tr><td>StAllc_THAdms (TH : 3)</td><td>16,3</td><td>57,6</td><td rowspan="2">-31,50</td><td></td></tr>
<tr><td>StAllc_THAdms (TH : 5)</td><td>22,6</td><td>57,4</td><td rowspan="2"></td><td rowspan="2">-51,00</td></tr>
<tr><td>StAllc_THAdms (TH : 10)</td><td>32,8</td><td>57,2</td><td></td></tr>
</table>

The last finding is related to StAllc_ARAdms policy. Although this policy is located efficient frontier borders in many scenarios, its performance is quite uncertain and change depending on the environment. For situation where traffic intensity is low, one-to-one matching (where some $x_{t,d}$ takes the value of 1) would be possible for StAllc policy optimization model like scenario 6 or scenario 12. In other words, optimal truck allocation ratios $x_{t,d}$ takes the value of one because each dock has enough capacity to accept all of the trucks from a specific truck type that have minimum penalty rate for that dock. It means that each specific truck type t should be directed to a determined dock d with probability equals to 1. Therefore, using StAllc_ARAdms policy results in same output with StAllc policy and the more $x_{t,d}$ values gets 1, the worse the effectiveness of the rule gets. To sum up, the efficient policies in different traffic intensity rates can be plotted as in Table 5.14

Table 5.14. Efficient policies in different traffic intensity rates

Environmet	State Independent (Static Policies)	Single Queue Policies	High TH Policies	Low TH Policies	StAllc_ ARAdnms Policy
High Traffic	Efficient	Efficient	Efficient	Not Efficient	Efficient
Low Traffic	Not Efficient	Efficient	Not Efficient	Efficient	Not Efficient

5.3.2. Correlation Between Expected Service Time And Penalty Rate

Impact of the correlation between expected service times and penalty rates is also investigated during the simulation phase. Positive level of correlation between service time and penalty rate indicates that some specific docks are preferable for some specific types of trucks because both their service times and penalty rates are low. On the other hand, in the negative correlation cases, there is a tradeoff between speed and penalty. Unloading docks that have high (low) expected unloading times have low (high) penalty rates for a specific truck type. Therefore, these docks are not preferable (preferable) for waiting time minimization (penalty minimization) aspect. Therefore, decision algorithms become much more important in such conflicting situations.

Waiting time or penalty centric rules (Min_UT, Min_P or Min _CbPr) give similar outputs in positively correlated environments. The reason is that ordinality of both service times, penalty rates are quite similar and any rule that minimizes average waiting times (penalties) are quite effective also to minimize average penalties (waiting times). Therefore any routing decision can take into account both the service time and penalty rate minimization. As seen in trade-off curves in Figure 5.7 and Figure 5.10, Min_UT, Min_P and Min_CbPr policy outputs concentrate in same area. Also as an numerical example, Table 5.15 shows that the waiting time difference between results of these policies is very low (about 10%) for positively correlated environment.

Table 5.15. Results of policies 2,3,4 in positively and negatively correlated environments

Av. Waiting Time (min.)	**Positive Correlation_**High Dispersion			**Negative Correlation_**High Dispersion		
	Low Traffic	Normal Traffic	High Traffic	Low Traffic	Normal Traffic	High Traffic
Min_UT	3,75	5,34	9,22	3,75	5,34	9,22
Min_P	3,69	5,21	8,33	7,83	12,5	30
Min _CbPr	3,5	4,99	8,13	5,33	7,25	14,3
Maximum (min)	3,75	5,34	9,22	7,83	12,5	30
Minimum (min)	3,5	4,99	8,13	3,75	5,34	9,22
% Delta	**7%**	**7%**	**12%**	**52%**	**57%**	**69%**

On the other hand, waiting time (Min_UT) and penalty centric rules (Min_P) results in different and contradictory results in terms of average waiting times and penalty rates where the correlation level between service times and penalty rates are 0 or negative. Tradeoff between the two important performance measures reveals in these environments. As expected, minimum waiting time rule concentrates on decreasing average waiting time at the expense of excessive level of penalty rates and vice versa for minimum penalty rule. In such a situation, using mixed-balancing hybrid rules (like Min_CbPr, StAllc_THAdms or StAllc_ARAdms) is very effective approach to keep both waiting times and penalty rates in acceptable limits. These rules can balance the levels of waiting times and penalty rates. Moreover, by using different threshold values, decision maker can also regulate the weighted importance of conflictive performance measures. For these reasons, it is much more effective to use mixed-balancing hybrid rules in negative correlation environments to arrange relative importance of conflicting performance measures. Investigating Figure 5.4 and Table 5.16 it is observed that Min_UT policy (time centric) and Min_P policy (penalty centric) give extreme outputs from two different perspective. However, using Min_CbPr (mixed-balancing) rule finds a balance point between these two edges. Actually, StAllc_THAdms policies are also balancing policies and system organizer can arrange the threshold values according to the weighted importance of two different performance measures.

Table 5.16. Results of policies 2,3,4 in negatively correlated environments

Policies	Negative Correlation_High Dispersion					
	Av. Waiting Time (min.)			Average Penalty Rate (m)		
	Low Traffic	Normal Traffic	High Traffic	Low Traffic	Normal Traffic	High Traffic
Min_UT	3,75	5,34	9,22	66,4	64,5	64,1
Min_P	7,83	12,5	30	35,3	35,3	32,3
Min _CbPr	5,33	7,25	14,3	49	49,3	48,1

Hence, if the service times and penalty rates are positively correlated, using one of the waiting time or penalty centric rules is logical because they are simpler to use and all these rules give the very similar results. On the other hand, for the environments where the correlation between penalty rates and waiting times are negative, minimum penalty and minimum service time rules give results that minimize one performance measure at the cost of another. Therefore, mixed rules are seen as much more logical to implement to get desired results for both performance measures.

StAllc policy and StAllc_THAdms policies with high level of threshold values are not preferable in the environments where correlation between service times and penalty rates are strictly negative. Since the allocation ratios ($x_{t,d}$) are determined to minimize expected penalty rates, these rules prioritize penalty minimization aspect. Therefore, emerging waiting times are very long especially in negatively correlated environments.. In other words, waiting times are almost maximum while penalty rates are minimized with static rule in a negatively correlated environment. Therefore, these rules are not implementable in real life situations. They can only give lower bounds for the average penalty rates in negatively correlated environments. To sum up, the efficient policies in different correlation levels can be plotted as in Table 5.17.

Table 5.17. Efficient policies in different correlation levels

Environment	Waiting Time-Centric Policies	Penalty Centric Policies	Mix-Balance Policies	State Independent (Static Policies)	Single Queue Policies	High TH Policies	Low TH Policies
Positive Correlation	Efficient	Efficient	Efficient	Efficient	Efficient	Efficient	Efficient
Negative Correlation	Not Efficient	Not Efficient	Efficient	Not Efficient	Efficient	Not Efficient	Efficient

5.3.3. Dispersion Of Penalty Rates

Dispersion of penalty rates indicates the difference between penalties of different truck type- unloading dock pairs. If the dispersion ratio is low, penalty rates are gathered and all the values are close to each other. On the other hand, penalty values are far from each other in highly dispersed environments.

Comparing high dispersion and low dispersion environments, it is observed that penalty centric policies (like Min_UT policy) cannot improve average penalty rates significantly in the environments where dispersion of penalty rates are low. Focusing on these kinds of rules may worsen average waiting times without meaningful improvement in penalty rates. Only small improvements in penalty rates can be achieved at the cost of huge level of increase in waiting times. This result is clearly seen in Figure 5.9 where the graph line in low dispersion environment is far steeper. Therefore, using waiting time centric policies are more effective in low dispersion environments.

On the other hand, penalty centric policies (like Min_P policy, StAllc or StAllc_THAdms with high thresholds) come into prominence when dispersion of penalty rates are huge because any deviation from ideal unloading dock to another dock results in high level of gap in terms of penalty rates. Therefore, very high level of penalty rates can be avoided by using penalty centric rules.

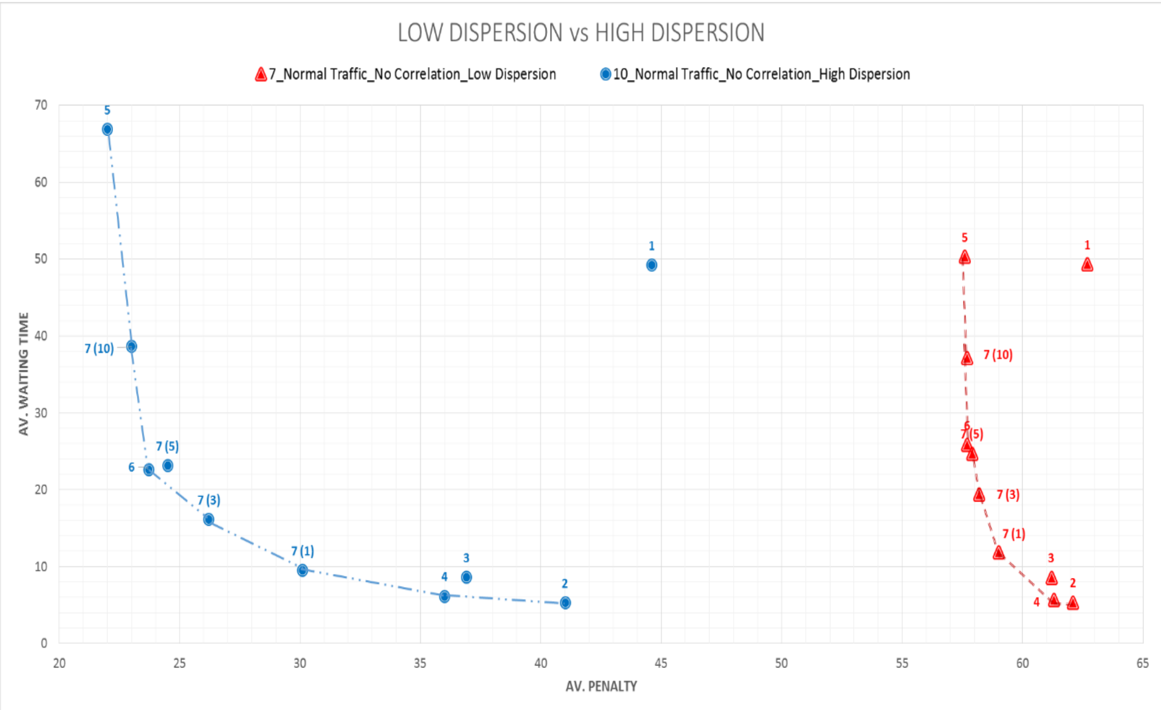

Figure 5.9. Comparison of the outputs of low dispersion vs high dispersion of penalty rates in uncorrelated and normal traffic intense environment

5.3.4. Utilization Rates Of Unloading Docks

Although achieving homogenous utilization rates for different unloading docks is not the main objective of the study, obtained results can also be evaluated in terms of utilization levels of docks.

Simulation results show that dock dedicated policies that do not allow admission from another dock's queue [Rand, StAllc or StAllc_THAdms with high threshold policies] cannot satisfy the workload balance among unloading docks. Since the static optimization model suggests the assignment rates to allocate incoming trucks to the docks that have lower penalty rates, some docks (with lower penalty rates) are consistently busy and others are idle. Especially for normal and high traffic intensity rates, the difference between utilization rates of unloading docks grows considerably as seen in Figure 5.10 and Figure 5.11. Actually, as a general approach, policies without dynamic admission (from other queues) opportunity are not so effective in terms of utilization balance of servers. These kinds of rules do not allow trucks to be served by another dock even if this dock is idle. Therefore, these kinds of rules are not preferable for the utilization balance of unloading docks.

Other rules that have single queues or that allow dynamic admission (from other dock-dedicated queues) are quite consistent to balance utilization levels of unloading docks. To illustrate, the utilization difference between unloading docks for Min_UT, Min_P, Min_CbPr or StAllc_THAdms(1) policies do not exceeds 5% as seen in Figure 5.10 and Figure 5.11

Observing the abovementioned characteristics, if balanced workload allocation among unloading docs are important, single queue policies or policies that allow dynamic admission (from other dock-dedicated queues) should be used instead of static and high-value threshold policies.

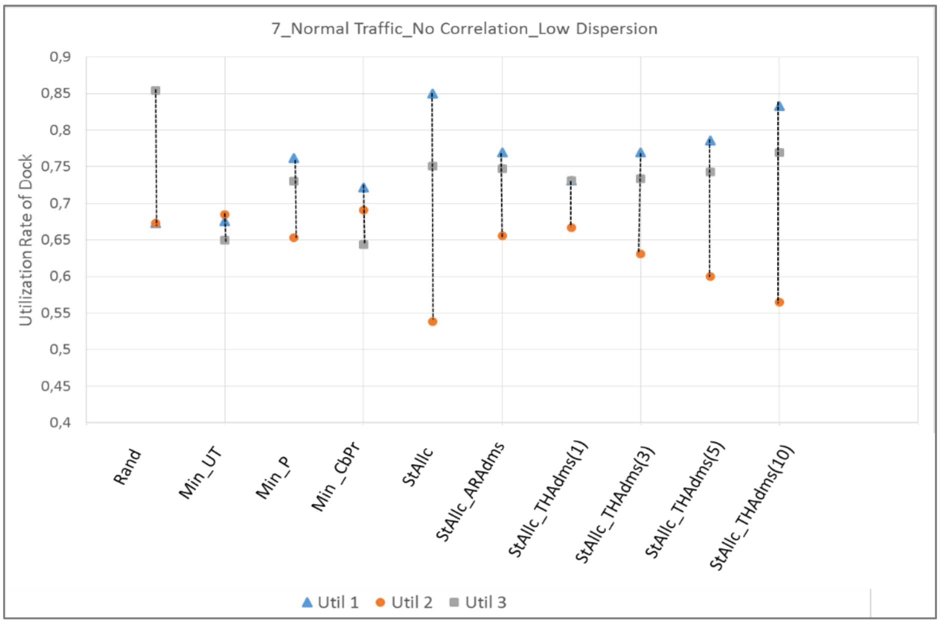

Figure 5.10. Difference in utilization levels of servers for different policies in normal traffic, no correlation and low dispersion environment

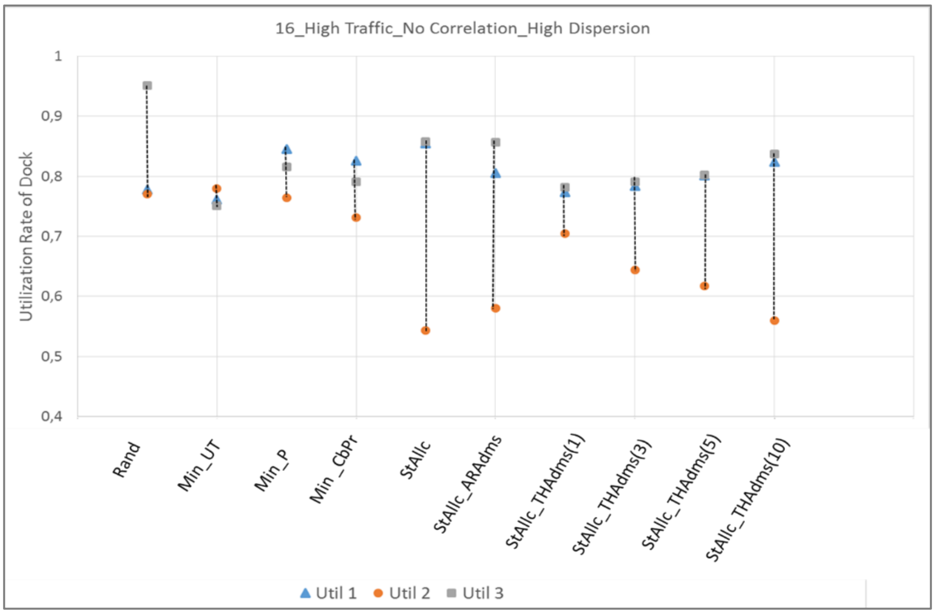

Figure 5.11. Difference in utilization levels of servers for different policies in normal traffic, no correlation and high dispersion environment

5.3.5. Inter-Arrival and Service Times Distribution Variance Effect on Outputs

Performance analysis of different routing policies are conducted under the assumption of exponentially distributed inter-arrival and service times. Due to the characteristics of exponential distribution, variance of the random variables are always equals to mean square value. That is:

$$\text{Variance} = V[x] = \frac{1}{\mu^2} \tag{5.12}$$

$$\text{Mean Square} = (E[x])^2 = \frac{1}{\mu^2} \tag{5.13}$$

For exponential distribution:

$$\frac{\text{Variance}}{\text{Mean Square}} = \frac{V[x]}{(E[x])^2} = 1 \tag{5.14}$$

To observe the effects of the variation of inter-arrival and service time distributions, base simulation modes are also run with high variance and low variance distributions. Scenario 7 with normal traffic intensity, no correlation between unloading times & penalty rates and low dispersion of penalty rates is used as the base model to compare effects of the distributions with different level of variations.

For this purpose, exponential distribution, lognormal distribution with variance to mean square ratio equals to 0.5 and lognormal distribution with variance to mean square ratio equals to 1.5 are used. Mean inter-arrival and service time values are kept as in Scenario 7 and variances are changed to satisfy mentioned variance to mean square ratio as seen in Table 5.18 Detailed results for different distribution variance scenarios can be seen in Appendix C.

Table 5.18. Different scenarios with high and low variance inter-arrival and service time distributions

Inter-Arrival Time		**Service Time**	
Distribution	Variance / (Mean Square)	Distribution	Variance / (Mean Square)
Exponential	1	Lognormal	0.5
Lognormal	0.5	Exponential	1
Lognormal	0.5	Lognormal	0.5
Exponential	1	Lognormal	1.5
Lognormal	1.5	Exponential	1
Lognormal	1.5	Lognormal	1.5
Exponential	1	Exponential	1

By simulating the scenarios given in Table 5.18, average waiting times and average penalty rates are obtained. The results are shown in Figure 5.12 and Figure 5.13. Comparing the response of different routing policies for different level of variances, two important conclusions can be drawn:

- Investigating the results for both average waiting time and average penalty aspects, it can be observed that the relative performance levels of the routing policies do not change with the different distribution variance scenarios. To illustrate, St_Allc policy gives the minimum penalty rate for all scenarios while minimum unloading time policy is the best to minimize average waiting times. Therefore, performance ranking of the policies are quite stable in different variance environments.

- While penalty rates are quite similar in different variance environments, average waiting times are fluctuated. Although this fluctuation is noticeable for high variance scenarios, it is maximum for static rules with dock-dedicated queue setting (St_Allc & Rand). The main reason is that static policies cannot tolerate the instant fluctuations because admission from different docks' queues are not allowed for these kinds of policies.

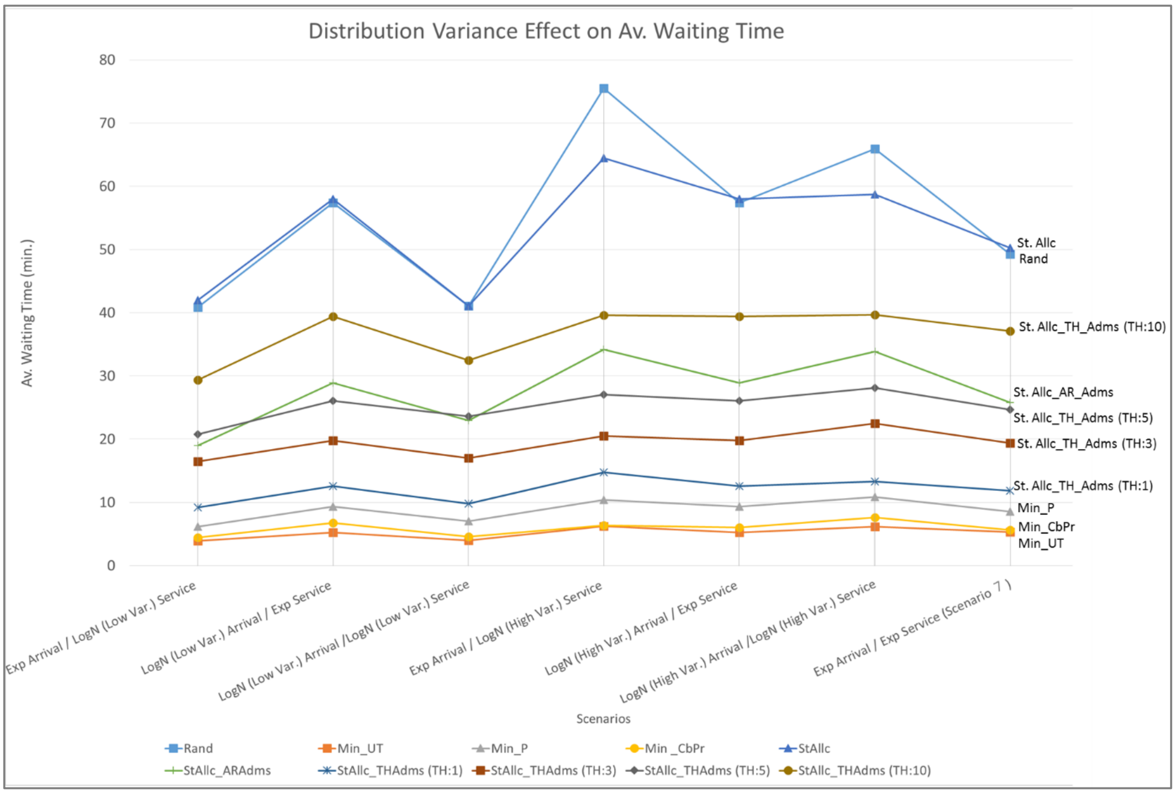

Figure 5.12. Distribution variance effect on average waiting times

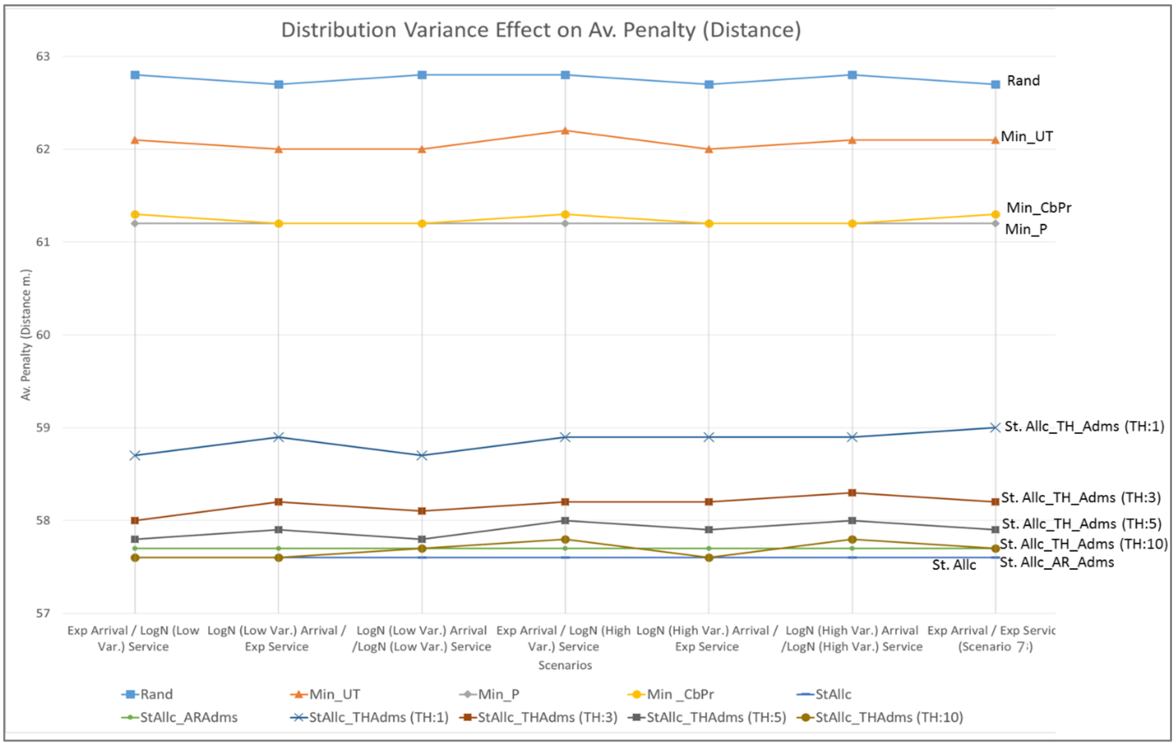

Figure 5.13. Distribution variance effect on average penalty rates

These general observations show that with the change in variance levels of inter-arrival and service time distributions, almost all of the policies keep their relative performance rankings stable. However, static or high threshold policies efficiency may change depending on the distribution variances. Other policies are quite robust to implement for different environments with varied distribution variance levels.

6. CASE STUDY: AN AUTOMOTIVE FACTORY

In this chapter, a real life problem from an automotive factory is investigated. As explained in Chapter 2.1, the case study involves unloading operations in the assembly warehouse of the automotive factory. There are three different types of materials called Domestic, Import and JIS –Just in Sequence- and coming trucks carry materials belong to one of these classes. Truck unloading operations are carried at three different docks and expected unloading times change for different truck type-dock pairs.

After unloading operations, parts are carried on to the shelves inside the warehouse area. Warehouse area is separated into three sub-sections for different classes of materials. Distance between the unloading docks and the sub-sections of warehouse to where the unloaded material should be carried results in some amount of penalty. Penalty rate per truck represents the distance and they are measured as meter per truckload. Therefore, dock selection for the unloading operation should aim to minimize this penalty. Additionally, this selection rule should also minimize the waiting time of coming trucks that passes until the unloading activity. Schematic representation of explained material flow is shown in Figure 3.1.

In the current situation, there is no routing policy for the coming trucks. Arriving trucks are directed to the docks from the main gate of the factory and they are waiting for unloading in front of the unloading docks. Since there is no routing policy, sometimes trucks may also be called by different idle docks due to several reasons. Considering the real life operations in the factory, it is seen hat approximately 250 trucks are coming to the warehouse in a day and average waiting time per truck is about one hour.

Referring to routing policies and conclusions of the simulation studies in Chapter 5, the case is investigated deeply and a solution model is proposed. Finally, the results of the case study is compared to the general results clarified in Chapter 5.5.5.

6.1. Input Analysis

Arena Input Analyzer Module is used to perform input analysis for the case study. By using fit menu, fitted distributions are determined and goodness of fit tests are performed. Corresponding square errors and p values are determined for both Chi-Square and Kolmogorov-Smirnov (K-S) tests. The results can be seen in Table 6.1 and Table 6.2 in detail. After these analyses, it is seen that exponential arrival and service times assumption fit the real data.

Data set includes 1-month period and about 4800 arrivals. As a result of distribution fitting tests, p values are obtained. Larger p-values implies that the suitability of the tested distribution in null hypothesis cannot be rejected. Therefore, for the p-values smaller than specified α, the null hypothesis can be rejected. For the greater p-values, it is concluded that the null hypothesis cannot be rejected and tested distribution can be fitted to the real data.

For the distribution testing procedure, null and alternative hypotheses are:

- H0: tested distribution sufficiently fits to the data
- H1: tested distribution is not a good fit

For the case study, hypothesis-testing procedure is conducted by specifying α significance level as 0.05. Then p-values of the tests are compared to pre-determined α=0.05 values. As a result of the tests it is concluded that the p-values are greater than 0.05 and fitted distribution cannot be rejected as shown in Table 6.1 and 6.2**.**

Table 6.1. Results of fitting distribution analysis for inter-arrival times

Inter-Arrival Times				
Truck Type	**Distribution**	**Square Error**	**p-Value (Chi Square)**	**p-Value (K-S)**
JIS	Expo (13.933)	0.000292	0.131	>0.15
Domestic	Expo (15.083)	0.000240	>0.75	>0.15
Import	Expo (26.667)	0.00116	0.136	>0.15

Table 6.2. Results of fitting distribution analysis for service times

Service Times				
Truck Type-Dock	**Distribution**	**Square Error**	**p-Value (Chi Square)**	**p-Value (K-S)**
1-1	Expo(1.783)+8.583	0.001259	0.0794	0.0354
1-2	Expo(1.850)+13.716	0.004066	0.0845	0.149
1-3	Expo(3.550)+13.600	0.004981	0.05	0.078
2-1	Expo(0.567)+11.567	0.001281	0.251	>0.15
2-2	Expo(3.100)+11.483	0.001215	0.234	>0.15
2-3	Expo(1.733)+12.833	0.001979	0.21	>0.15
3-1	Expo(2.033)+15.150	0.004187	0.141	>0.15
3-2	Expo(3.800)+10.450	0.002850	0.344	>0.15
3-3	Expo(1.383)+11.933	0.000621	0.192	>0.15

Penalty rates, on the other hand, are determined by measuring the distance between the unloading dock "d" and the center of the warehouse area of the material transported by truck *t*. By this way penalty rates are determined for each truck-unloading dock *(t,d)* pair as shown in Table 6.3

Table 6.3. Penalty Rates for case study

Penalty Rates (Distances)									
(Truck Type-Dock)	**1--1**	**1--2**	**1--3**	**2--1**	**2--2**	**2--3**	**3--1**	**3--2**	**3--3**
Distance (m)	75.6	136.8	217.8	127.8	50.4	113.4	239.4	140.4	57.6

6.2. Simulation Model

Simulation models explained in Chapter 4.3 are used by changing parameters according to the real life data. Inter-arrival times of trucks, penalty rates, service times and other parameters are changed according to the obtained values after input analysis explained in Chapter 6.1. Investigating the inputs of the case study, it can be asserted that the environment is similar to studied experimental scenario where traffic intensity is normal, correlation level between service times and penalty rates are positive and dispersion of penalty rates are high. From the generalized results given in Chapter 5, it can be foreseen that, StAllc_THAdms policies with low threshold values, Min_CbPr or Min_P policies will be efficient to use.

6.2.1. Warm-Up Period and Replication Number

Warm-Up period and replication number for each simulation run is determined as explained in Chapter 5.3.1 and in Chapter 5.3.2. As a result, used replication length, warm-up period and replication numbers of different simulation models are shown in Table 6.4, accordingly.

Table 6.4. Replication length, warm-up period and replication numbers for different policies in the case study

Policy	Rep. Length	Warm-Up	Repl. Num
Rand	120.000	30.000	9
Min_UT	100.000	25.000	5
Min_P	120.000	30.000	5
Min _CbPr	100.000	25.000	5
StAllc	160.000	40.000	7
StAllc_ARAdms	160.000	40.000	7
StAllc_THAdms(1)	100.000	25.000	7
StAllc_THAdms(3)	100.000	25.000	7
StAllc_THAdms(5)	120.000	30.000	7
StAllc_THAdms(10)	120.000	30.000	7

6.2.2. Verification and Validation

Verification process is conducted in the same way as explained in chapter 5.3.3 and it is observed that simulation models are consistent and they run as intended. The validation process, on the other hand, is needed to assure that the models reflect the real life situation and they are the accurate representation of the investigated system.

Validation tests for the general simulation models are conducted in Chapter 5.3.3 and it is seen that theoretical results gained by analytical formulas are drop into the confidence intervals obtained from simulation model outputs. In this chapter, the validation of the model is also supported by investigating the consistence of simulation outputs and real case data.

The first and important step to validate the simulation models is to create appropriate arrival and service events with consistent time intervals. To achieve the consistency of fitted distributions with real event times, statistical methods are used for input analysis as reported deeply in section 6.1. After simulation runs, it can be shown that average daily coming truck numbers in real life and created truck numbers in simulations are consistent with each other as seen in Table 6.5.

Table 6.5. Incoming trucks in real case and number of trucks created in simulation

Truck Number/day (average)					
Simulation Models (Av.)			Real Case (Av.)		
JIS	Domestic	Import	JIS	Domestic	Import
[94.3 , 110.9]	[74.9 , 89.9]	[42.1, 52.5]	105.6	82.4	47.3

Another approach is simulating current real life situation and comparing simulation outputs with real performance measures. Considering the current implementations in automotive factory, it is seen that there is not any systematic algorithm with a clever policy to allocate coming trucks to unloading docks. Moreover, some troubles may occur during daily unloading or warehousing activities. Despite all of these extreme situations, it is assumed that randomized policy can roughly projects the current situation. The comparison table for average waiting time and penalty (distance) rates in real life situation and in randomized policy simulation model is shown in Table 6.6. As seen in the table, simulation results and real life situation outputs are not far from each other and the outcome implies that the simulation model can be used as a representation of the actual situation.

Table 6.6. Comparison of realized and simulated performance measures

	Real Case W(q) (min)	**Simulation Output 95% CI W(q) (min)**
Waiting in Queue	52	[48.32 , 59.88]

Considering verification and validation analysis for experimental models in Chapter 5.2.3 and the analysis in this chapter, it can be concluded that the simulation models are run as intended and they represent the real-life situation reasonably. Therefore, the numerical results are significant for the appropriate reflection of the observed system.

6.3. Numerical Results

Simulation models for seven different policies are run with actual parameters of the real system. The results are revealed in Table 6.7 and Figure 6.1.

Table 6.7. Simulation result for the case study

Policy Name	Waiting Time	Half Width (0.95)	Av. Pen.	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
Rand	54,1	6,38	127	0,226	10,1	0.79	0,78	0,89
Min_UT	9,36	1,19	107	0,597	1,64	0,77	0,76	0,81
Min_P	9,91	0,694	96,8	0,273	1,74	0,76	0,79	0,77
Min_CbPr	9,73	0,843	96,7	0,568	1,71	0,76	0,79	0,77
StAllc	40,6	2,19	67,1	0,157	6,26	0,74	0,84	0,61
StAllc_ARAdms	20,3	1,08	69,2	0,0792	3,56	0,74	0,68	0,77
StAllc_THAdms(1)	12,5	1,5	79,8	0,452	2,2	0,75	0,78	0,73
StAllc_THAdms(3)	19,3	1,04	73,1	0,47	3,73	0,81	0,81	0,69
StAllc_THAdms(5)	26,1	1,22	70,4	0,344	4,81	0,83	0,83	0,64
StAllc_THAdms(10)	36,2	2,02	67,3	0,321	6,63	0,85	0,85	0,57

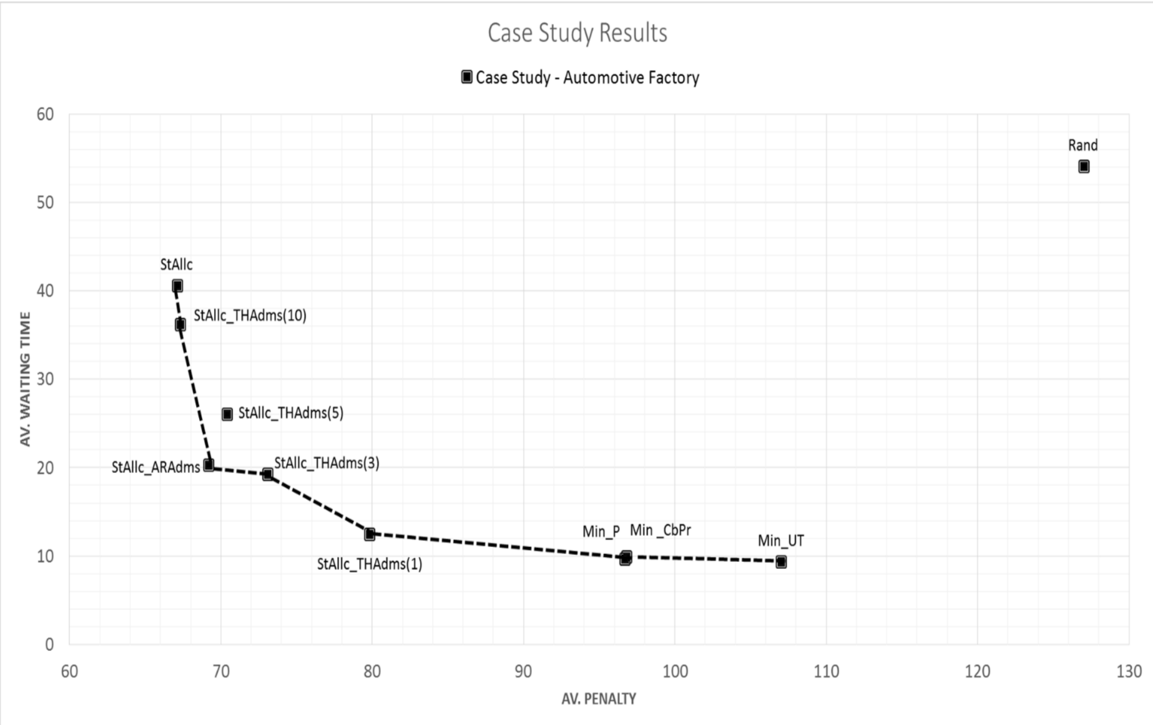

Figure 6.1. Graph for simulation result for the case study

The real life parameters are similar to the experimental environment where traffic intensity is medium, correlation level between service times and penalty rates are positive and dispersion of penalty rates are high. The results are observed as consistent with the generalized conclusions in Chapter 5.4. Indeed, as concluded in Chapter 5.4, Min_P, Min_CbPr, Min_UT policy and StAllc_THAdms policies with low threshold values give similar results due to positive correlation and they are located in the efficient frontier.

StAllc policy gives the minimum average penalty rate but the average waiting time is very large. As explained in Chapter 5, using this rule in real life activities is not logical due to excessive waiting times. Even if the penalty rate has higher importance comparing the waiting time, StAllc_THAdms policies with different threshold levels are more efficient to use.

StAllc_ARAdms policy is also located in efficient frontier line and gives quite good results in terms of both penalty and waiting time minimization aspects. However, it is robustness level is not reliable. The reason is that with different $x_{t,d}$ values, relative performance of StAllc_ARAdms policy can change pretty much as explained in Chapter 4.3.2.

Randomized rule, on the other hand, is strictly dominated by all other rules Therefore; it cannot be a candidate policy for the implementation.

Min_P policy is also dominated by Min_CbPr policy. Actually, in the environments where penalty rates and waiting times are positively correlated, Min_P, Min_UT and Min_CbPr policies give very similar results.

For StAllc_THAdms policies with threshold values greater than 3, it is observed that improving average penalty rate results in considerable increase in average waiting time. Additionally, StAllc_THAdms (5) policy is also dominated by some other policies. Such policies with high threshold values can be implemented in cases where penalty rate is much more important compared to waiting time. However, for the automotive company where the case study is conducted, it is not the case. For this company, average waiting time and average penalty rate are somehow equally important. For this reason, using StAllc_THAdms algorithms with larger threshold values is not an appropriate choice.

Hence, the summary of the results obtained from the case study can be explained as follows:

- “Rand”, “Min_P” and “StAllc_THAdms(5)” policies are strictly dominated by other policies.

- “StAllc_THAdms” policies with threshold values “greater than 3” are also not efficient for the studied case because the cost of small improvement in penalty rates is become very huge in terms of waiting times of trucks as the threshold value increases.

- “Min_UT”, “Min_CbPr”, “StAllc_ARAdms” “StAllc_THAdms(1)” and “StAllc_THAdms(3)” seem as candidate policies to implement in the system. Potential improvement rates with these proposed policies are shown in Table 6.9.

Decision maker can select one of these rules by evaluating relative tradeoffs between the main performance measures.

Table 6.8. Performance measures in current system

	Performance Measure	Actual Data
Current Situation	Waiting Time (min/truck)	55
	Penalty -Distance- (m/truck load)	145

Table 6.9. Proposed routing policies and corresponding improvement potentials

Proposed Rule	Performance Measure	Output	Improvement Potential
Min Time	Waiting Time (min/tr)	9,36	82,98%
	Penalty -Distance (m/tr load)	107	26,21%
Min Time&Pen.	Waiting Time (min/tr)	9,73	82,31%
	Penalty -Distance (m/tr load)	96,7	33,31%
Swap acc. TH:1	Waiting Time (min/tr)	12,1	78,00%
	Penalty -Distance (m/tr load)	79,6	45,10%
Swap acc. TH:3	Waiting Time (min/tr)	20,5	62,73%
	Penalty -Distance (m/tr load)	71,8	50,48%

7. CONCLUSION AND FUTURE RESEARCHES

In this thesis, several routing algorithms for the multi-class docking system of an automotive manufacturer are examined. The model mainly includes multi-type trucks served by multi-type docks that have different service times and penalty rate parameters. Average waiting time of a truck and average penalty rate occurred per truck are defined as main performance measures for the examined system. The fundamental decision is to match incoming trucks and unloading docks to optimize the defined performance measures. To achieve this goal and construct a well-defined policy, several routing policies are examined.

Different system environments are constructed to test the efficiency of the investigated routing policies. Traffic intensity rate, correlation between the input parameters (service time vs penalty rate) and dispersion of penalty rates are the factors that differentiate the environment in which routing policies operate. Different levels of these factors are set and 18 different environment setups are constructed. All the candidate routing policies are simulated in different experimental environments and important conclusions are drawn to help to understand the relative performance of the routing policies in different environments.

After analyzing the output, it is observed that while some queuing systems and routing policies are very efficient to minimize average waiting time of the trucks others are powerful to reduce average penalty rates. To balance these two measures, some hybrid policies are proposed and it is seen that these hybrid policies help the system owner to arrange relative importance of two different performance measures by changing threshold parameters.

In the last part of the study, a real problem from an automotive factory is examined. The implementation of different policies are simulated with actual data and in a real system setting. The outputs and obtained performance measures are compared with the experimental scenario simulations. It is observed that the result are consistent with the generalized results drawn in Chapter 5. Also some policies are suggested for the implementation at the unloading docks of the automotive factory. It is anticipated that average waiting time and penalty rate (distance) can be improved by up to 80% and 45% respectively.

This study examines the trade-off curves for average waiting time and average penalty rate due to the distance covered. A future extension to this study will be to quantify the exact unit cost of waiting time and penalty rate and then searching for an optimal routing policy that minimize the total cost of the system. Also, differentiating using differentiated unit (waiting) time costs for truck types may be another extension of the study.

In this study, inter-arrival and service time distributions are time-independent. However, learning characteristics of the servers or different arrival rates in a daytime can generate time dependent arrival and service patterns in some systems. Therefore, the inclusion of time dependent service and arrival processes can be a valuable extension for this study.

REFERENCES

1. Rodrigue, J. P. (2012). Supply chain management, logistics changes and the concept of friction. *Cities, Regions and Flows*, 2842-2843.

2. Newell, C. (2013). *Applications of queueing theory* (Vol. 4). Springer Science & Business Media.

3. De Véricourt, F., & Zhou, Y. P. (2005). Managing response time in a call-routing problem with service failure. *Operations Research*, *53*(6), 968-981.

4. Armony, M. (2005). Dynamic routing in large-scale service systems with heterogeneous servers. *Queueing Systems*, *51*(3), 287-329.

5. Bowersox, D. J., Closs, D. J., & Cooper, M. B. (2002). *Supply chain management* (Vol. 2). New York, NY: McGraw-Hill.

6. Wilson, R. (2013). *State of Logistics Report.* Council of Supply Chain Management Professionals (CSCMP).

7. Atar, R. (2008). Central limit theorem for a many-server queue with random service rates. *The Annals of Applied Probability*, *18*(4), 1548-1568.

8. Armony, M., & Ward, A. R. (2010). Fair dynamic routing in large-scale heterogeneous-server systems. *Operations Research*, *58*(3), 624-637.

9. Gurvich, I., & Whitt, W. (2009). Queue-and-idleness-ratio controls in many-server service systems. *Mathematics of Operations Research*, *34*(2), 363-396.

10. Zhan, D., & Ward, A. R. (2013). Threshold routing to trade off waiting and call resolution in call centers. *Manufacturing & Service Operations Management*, *16*(2), 220-237.

11. Mehrotra, V., Ross, K., Ryder, G., and Zhou, Y. P. (2012). Routing to manage resolution and waiting time in call centers with heterogeneous servers. *Manufacturing & service operations management*, *14*(1), 66-81.

12. Sophia, M.E, and Chiemeke, S.C. (2016). Performance Evaluation of Call Resolution Oriented Routing Rules to Enhance Resolution Rates. *International Journal of Computer Applications 143(8).*

13. M. van Tits and H. van der Veeken, "Simulation of a queueing problem with balking," *Simulation, vol. 35, pp. 88-93, September I, 1980.*

14. Kleinrock, L. and Finkelstein, R.P.: Time dependent priority queues. *Oper. Res. 15, 104–116 (1967).*

15. P. W. Glynn and D. L. Iglehart (1988). Simulation Methods for Queues: An Overview. Queueing Systems: *Theory and Applications, Vol. 3, 221-256.*

16. Jang, J., Chung, J., Suh, J. and Rhee, J., "Estimation of the mean waiting time of a customer subject to balking: a simulation study, " *International Journal of Flexible Manufacturing Systems, vol. 18, pp. 121-144,2006.*

17. Özlem, Ș. (2008), Simulation Of The Vessel Traffic In The Strait Of Istanbul, *Master Thesis.*

18. Woottichaiwat. S, (2015). Efficiency Improvement of Truck Queuing System in the Freight Unloading Process Case Study of a Private Port in Songkhla Province. *Ilpakorn University Science And Technology Journal, Vol 9, No 2.*

19. Gopakumar, B., Sundaram, S., Shengyong, W., Koli, S. and Srihari, K. (2008), "Simulation based approach for dock allocation in a food distribution center", *Proceedings of the 2008 Winter Simulation Conference, Miami, FL, USA, pp. 2750-5.*

20. Iannoni, A.P. and Morabito, R., 2006. A discrete simulation analysis of a logistics supply system. *Transport Research Part E 42, 191–210.*

21. C-Y. Liong and C.S.E. Loo, ''‘A simulation study of warehouse loading and unloading systems using Arena’' *Journal of Quality Measurement and Analysis: JQMA (2009).*

22. Victoria Transport Policy Institute (VTPI), *Transportation Demand Management (TDM) Encyclopedia*, 2017. http://www.vtpi.org/tdm/tdm66.htm, accessed at June 2017.

23. Law, A.M. and Kelton, W.D. Simulation modeling & analysis, *3rd edn. McGraw-Hill, Inc., New York,* (2000).

24. Figueiredo Filho, D. B., Paranhos, R., Rocha, E. C. D., Batista, M., Silva Jr, J. A. D., Santos, M. L. W. D. and Marino, J. G. (2013) 'When is statistical significance not significant?', *Brazilian Political Science Review, 7(1), p 31-55.*

25. Montgomery, D.C., *Design and Analysis of Experiments*, 5th ed., John Wiley & Sons, New York, 2001.

APPENDIX A: INPUT PARAMETERS OF EXMERIMENTAL SCENARIOS

Table A.1. Service times in different scenarios

Service Times (All Scenarios Same)	
Truck Type	**Distribution**
1-1	Expo(9)
1-2	Expo(15)
1-3	Expo(18)
2-1	Expo(12)
2-2	Expo(12)
2-3	Expo(15)
3-1	Expo(18)
3-2	Expo(15)
3-3	Expo(12)

Table A.2. Inter-arrival times in different scenarios

Inter-Arrival Times				
Truck Type	**Traffic Intensity: Very High**	**Traffic Intensity: High**	**Traffic Intensity: Medium**	**Traffic Intensity: Low**
JIS	Expo(10)	Expo(12)	Expo(12)	Expo(15)
Domestic	Expo(15)	Expo(15)	Expo(20)	Expo(22.5)
Import	Expo(20)	Expo(30)	Expo(35)	Expo(30)

Table A.3. Penalty rates in different scenarios

Truck Type-Dock	Correlation Level : Positive Dispersion: High	Correlation Level : Positive Dispersion: Low	Correlation Level : No Corr. Dispersion: High	Correlation Level : No Corr. Dispersion: Low	Correlation Level : Negative Dispersion: High	Correlation Level : Negative Dispersion: Low
1-1	10	49	19	54	95	66
1-2	92	56	13	71	44	53
1-3	70	63	34	45	12	56
2-1	40	62	94	56	33	65
2-2	10	55	44	81	88	62
2-3	96	63	90	75	62	48
3-1	95	67	21	64	11	50
3-2	40	61	73	69	10	58
3-3	10	49	16	53	98	61

APPENDIX B: OUTPUTS OF EXPERIMENTAL SCENARIOS

Table B.1. Results of scenario 1

	1_Low Traffic_No Correlation_Low Dispersion							
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	29,8	2,27	62,7	0,0735	4,3	0,63	0,599	0,752
2	3,75	0,411	62,2	0,114	0,544	0,597	0,616	0,582
3	5,76	0,369	60,9	0,126	0,838	0,714	0,574	0,663
4	4,12	0,272	61,3	0,144	0,596	0,672	0,634	0,566
5	50,3	8,26	57,1	0,0403	7,26	0,851	0,463	0,667
6	22,5	1,4	57,2	0,0719	3,24	0,764	0,576	0,668
7 (1)	9,7	0.394	58,6	0,108	1,4	0,698	0,595	0,669
7 (3)	16,3	0,842	57,6	0,0682	2,34	0,752	0,542	0,668
7 (5)	22,6	1,48	57,4	0,0882	3,26	0,785	0,518	0,679
7 (10)	32,8	2,18	57,2	0,073	4,73	0,814	0,484	0,68

Table B.2. Results of scenario 2

2_Low Traffic_Negative Correlation_Low Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	29,8	2,27	58,3	0,042	4,3	0,63	0,599	0,752
2	3,75	0,411	60,7	0,0699	0,544	0,597	0,616	0,582
3	8,43	0,537	55,9	0,044	1,22	0,68	0,678	0,742
4	5,19	0,155	57,6	0,0598	0,751	0,693	0,629	0,612
5	25,7	0,945	54,8	0,0555	2,28	0,57	0,564	0,568
6	8,55	0,62	55,1	0,0542	0,803	0,467	0,665	0,583
7 (1)	6,86	0,29	59,9	0,0435	0,991	0,585	0,575	0,598
7 (3)	11,1	0,401	55,1	0,0528	1,6	0,574	0,57	0,568
7 (5)	13,6	0,701	54,9	0,0875	1,96	0,564	0,576	0,568
7 (10)	19,4	0,491	54,9	0,0373	2,23	0,573	0,572	0,564

Table B.3. Results of scenario 3

3_Low Traffic_Positive Correlation_Low Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	29,8	2,27	58,5	0,0539	4,3	0,63	0,599	0,752
2	3,75	0,441	55,4	0,105	0,544	0,597	0,616	0,582
3	3,65	0,214	55,4	0,0671	0,529	0,62	0,614	0,57
4	3,54	0,191	55,3	0,096	0,514	0,609	0,62	0,596
5	12,5	1,01	50,9	0,0252	1,81	0,603	0,54	0,398
6	12,9	1,41	50,9	0,0348	1,87	0,603	0,546	0,4
7 (1)	5,7	0,211	52,6	0,0869	0,82	0,579	0,56	0,492
7 (3)	9,12	0,479	51,3	0,0745	1,32	0,583	0,542	0,439
7 (5)	10,6	0,139	51	0,504	1,52	0,589	0,532	0,403
7 (10)	12,1	0,585	50,9	0,0241	1,74	0,594	0,544	0,393

Table B.4. Results of scenario 4

4_Low Traffic_No Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	29,8	2,27	44,7	0,161	4,3	0,63	0,599	0,752
2	3,75	0,411	40,1	0,411	0,544	0,597	0,616	0,582
3	5,03	0,303	36,2	0,189	0,727	0,707	0,546	0,655
4	4,16	0,37	35,7	0,235	0,326	0,999	0	0,934
5	37,00	4,31	29,80	0,14	5,36	0,36	0,86	0,40
6	16,10	0,73	24,40	0,22	1,52	0,46	0,73	0,41
7 (1)	6,39	0,189	35,5	0,134	0,925	0,507	0,666	0,516
7 (3)	10,9	0,252	26	0,29	1,56	0,428	0,733	0,46
7 (5)	15	0,569	27,2	0,2	2,16	0,4	0,781	0,429
7 (10)	23,6	1,72	23,6	0,325	3,41	0,375	0,826	0,425

Table B.5. Results of scenario 5

5_Low Traffic_Negative Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	29,8	2,27	49,1	0,242	4,3	0,63	0,599	0,752
2	3,75	0,441	66,4	0,555	0,544	0,597	0,616	0,582
3	7,83	0,604	35,3	0,324	1,13	0,686	0,636	0,773
4	5,33	0,223	49	0,211	0,772	0,712	0,534	0,705
5	52,80	5,30	38,10	0,06	7,67	0,61	0,68	0,86
6	28,8	1,6	19,3	0,088	4,15	0,602	0,659	0,868
7 (1)	11,5	1,14	29,2	0,32	1,66	0,666	0,667	0,725
7 (3)	19,6	0,28	23,6	0,183	2,84	0,646	0,683	0,774
7 (5)	25,1	0,567	21,6	0,134	3,63	0,625	0,683	0,791
7 (10)	36,3	1,82	20,2	0,137	5,26	0,611	0,695	0,818

Table B.6. Results of scenario 6

6_Low Traffic_Positive Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	29,8	2,27	50,5	0,193	4,3	0,63	0,599	0,752
2	3,75	0,411	33,7	0,554	0,544	0,597	0,616	0,583
3	3,69	0,243	31,9	0,254	0,535	0,607	0,601	0,594
4	3,5	0,522	32,5	0,203	0,506	0,582	0,613	0,587
5	12,5	1,32	10	0	1,81	0,596	0,546	0,392
6	12,5	1,32	10	0	1,81	0,596	0,546	0,392
7 (1)	5,55	0,0776	18,5	0,196	0,801	0,562	0,559	0,505
7 (3)	8,67	0,261	11,9	0,126	1,25	0,581	0,536	0,426
7 (5)	10,9	0,336	10,6	0,0892	1,57	0,593	0,536	0,41
7 (10)	12	0,534	10	0,0194	1,73	0,595	0,537	0,392

Table B.7. Results of scenario 7

7_Normal Traffic_No Correlation_Low Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	49,3	1,49	62,7	0,0806	7,96	0,673	0,673	0,854
2	5,34	0,534	62,1	0,0731	0,868	0,675	0,685	0,65
3	8,57	1,11	61,2	0,0725	1,39	0,762	0,653	0,73
4	5,67	0,572	61,3	0,0752	0,916	0,722	0,691	0,644
5	50,30	5,32	57,60	0,03	8,13	0,85	0,54	0,75
6	25,8	2,35	57,7	0,048	4,18	0,77	0,656	0,747
7 (1)	11,9	2,7	59	0,857	1,92	0,731	0,667	0,731
7 (3)	19,4	1,85	58,2	0,0984	3,13	0,77	0,631	0,734
7 (5)	24,7	0,735	57,9	0,113	3,98	0,786	0,6	0,743
7 (10)	37,1	4,26	57,7	0,0778	6,2	0,833	0,565	0,77

Table B.8. Results of scenario 8

8_Normal Traffic_Negative Correlation_Low Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	49,3	1,49	58,6	0,505	7,96	0,673	0,673	0,854
2	5,34	0,534	61	0,0523	0,868	0,675	0,685	0,65
3	14,4	1,31	56,4	0,0343	2,33	0,752	0,751	0,816
4	5,19	0,155	57,6	0,0598	0,751	0,693	0,629	0,612
5	72,20	8,40	53,90	0,03	12,50	0,76	0,86	0,85
6	28,9	3,11	54,6	0,0966	4,68	0,858	0,792	0,736
7 (1)	18,8	1,94	55,2	0,0667	3,05	0,767	0,784	0,794
7 (3)	29,6	2,32	54,4	0,0927	4,79	0,776	0,817	0,824
7 (5)	38,4	2,21	54,2	0,0659	6,25	0,779	0,834	0,829
7 (10)	38,4	2,21	54,2	0,0659	6,25	0,779	0,834	0,829

Table B.9. Results of scenario 9

9_Normal Traffic_Positive Correlation_Low Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	49,3	1,49	58,5	0,0607	7,96	0,673	0,673	0,854
2	5,34	0,534	55,6	0,0842	0,868	0,675	0,685	0,65
3	4,93	0,21	55,6	0,0723	0,796	0,683	0,674	0,631
4	5,35	0,206	55,5	0,0356	0,869	0,684	0,684	0,634
5	20,5	1,67	50,9	0,00858	3,32	0,75	0,599	0,341
6	21,7	2,27	50,8	0,0159	3,53	0,756	0,608	0,344
7 (1)	7,15	0,447	53	0,0869	1,15	0,668	0,628	0,53
7 (3)	11,2	0,58	51,7	0,0688	1,81	0,696	0,613	0,432
7 (5)	14	0,262	51,2	0,0427	2,27	0,718	0,608	0,382
7 (10)	17,9	0,845	50,9	0,0608	2,9	0,743	0,599	0,352

Table B.10. Results of scenario 10

10_Normal Traffic_No Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	49,3	1,49	44,6	0,196	7,96	0,673	0,673	0,854
2	5,34	0,534	41	0,315	0,868	0,675	0,685	0,65
3	8,66	0,383	36,9	0,112	1,41	0,772	0,645	0,721
4	6,2	0,426	36	0,259	1,01	0,758	0,624	0,711
5	66,90	11,40	22,00	0,09	10,90	0,85	0,32	0,85
6	22,7	0,708	23,7	0,0787	2,6	0,736	0,42	0,814
7 (1)	9,58	0,572	30,1	0,192	1,55	0,72	0,564	0,734
7 (3)	16,2	1,9	26,2	0,308	2,63	0,762	0,469	0,767
7 (5)	23,2	0,605	24,5	0,247	3,77	0,802	0,419	0,794
7 (10)	38,7	2,83	23	0,114	6,31	0,831	0,362	0,838

Table B.11. Results of scenario 11

11_Normal Traffic_Negative Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	49,3	1,49	48,7	1,49	7,96	0,673	0,673	0,854
2	5,34	0,534	64,5	0,418	0,868	0,675	0,685	0,65
3	12,5	1,03	35,3	0,256	2,03	0,752	0,728	0,837
4	7,25	0,161	49,3	0,169	1,17	0,762	0,619	0,762
5	64,6	6,77	38,7	0,0904	10,5	0,65	0,849	0,85
6	33,3	0,995	23,4	0,13	3,76	0,797	0,734	0,851
7 (1)	15,9	1,48	31,2	0,182	2,59	0,736	0,767	0,778
7 (3)	24	0,998	26,5	0,414	3,9	0,711	0,795	0,8
7 (5)	31,3	1,88	24,7	0,0403	5,08	0,69	0,808	0,819
7 (10)	45,2	2,96	23,3	0,107	7,32	0,655	0,833	0,835

Table B.12. Results of scenario 12

12_Normal Traffic_Positive Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	49,3	1,49	49,8	1,49	7,96	0,673	0,673	0,854
2	5,34	0,534	33,9	0,402	0,868	0,675	0,685	0,65
3	5,21	0,497	32	0,296	0,846	0,673	0,673	0,654
4	4,99	0,335	32,5	0,24	0,809	0,658	0,679	0,651
5	20,7	1,79	10	0	3,35	0,756	0,596	0,341
6	20,7	1,79	10	0	3,35	0,756	0,596	0,341
7 (1)	6,82	0,617	19,9	0,493	1,1	0,654	0,627	0,537
7 (3)	11,3	0,559	13,6	0,372	1,82	0,696	0,615	0,436
7 (5)	14	0,844	11,5	0,179	2,27	0,721	0,605	0,382
7 (10)	18,3	0,865	10,3	0,119	2,96	0,744	0,6	0,357

Table B.13. Results of scenario 13

13_High Traffic_No Correlation_Low Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	130	22,9	63,7	0,072	23,7	0,778	0,771	0,951
2	9,22	0,837	63	0,0965	1,7	0,761	0,78	0,751
3	17,2	2,02	62,4	0,0508	3,15	0,844	0,771	0,827
4	10,4	0,781	62	0,106	1,92	0,798	0,786	0,751
5	77,80	8,81	58,80	0,03	14,30	0,85	0,85	0,85
6	33,70	3,62	59,20	0,11	6,18	0,92	0,86	8,72
7 (1)	23,5	1,59	59,9	0,053	4,31	0,829	0,833	0,839
7 (3)	31,7	2,13	59,3	0,0959	5,8	0,837	0,843	0,841
7 (5)	41,9	4,05	59,1	0,106	7,68	0,843	0,848	0,854
7 (10)	58,3	5,66	58,9	0,0886	10,7	0,849	0,847	0,854

Table B.14. Results of scenario 14

14_High Traffic_Negative Correlation_Low Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	130	22,9	58,4	0,0308	23,7	0,778	0,771	0,951
2	9,22	0,837	60,8	0,0428	1,7	0,761	0,78	0,751
3	43,4	4,11	55,9	0,058	7,99	0,887	0,881	0,91
4	16,5	1,02	57,8	0,0546	3,03	0,845	0,799	0,806
5	78,90	7,47	55,80	0,04	14,50	0,86	0,85	0,85
6	38,90	2,73	56,40	0,83	5,28	0,85	0,82	0,86
7 (1)	25,4	2,32	56,5	0,0578	4,65	0,837	0,834	0,85
7 (3)	33,9	2,17	56,1	0,0492	6,21	0,839	0,839	0,85
7 (5)	41,3	1,72	56	0,0429	7,56	0,84	0,847	0,849
7 (10)	55,55	3,63	55,8	0,0495	10,2	0,842	0,846	0,843

Table B.15. Results of scenario 15

15_High Traffic_Positive Correlation_Low Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	130	22,9	58,4	0,0337	23,7	0,778	0,771	0,951
2	9,22	0,837	55,6	0,169	1,7	0,761	0,78	0,751
3	8,08	0,559	55,4	0,0765	1,48	0,766	0,769	0,734
4	8,43	1,01	55,3	0,0727	1,55	0,764	0,771	0,73
5	31	2,46	51,2	0,0256	5,53	0,755	0,795	0,401
6	30	2,17	51,2	0,0147	5,49	0,754	0,795	0,396
7 (1)	9,86	0,428	53,2	0,757	1,8	0,723	0,737	0,634
7 (3)	14	0,588	52	0,078	2,56	0,725	0,745	0,531
7 (5)	18,5	0,875	51,6	0,0728	3,39	0,736	0,768	0,475
7 (10)	24,5	1,54	51,3	0,0471	4,49	0,747	0,78	0,416

Table B.16. Results of scenario 16

16_High Traffic_No Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	130	22,9	44,5	0,184	23,7	0,778	0,771	0,951
2	9,22	0,837	41,1	0,24	1,7	0,761	0,78	0,751
3	15,5	1,02	36,4	0,127	2,85	0,846	0,764	0,816
4	9,81	0,377	35,3	0,161	1,8	0,826	0,731	0,791
5	62,00	5,60	24,10	0,13	11,40	0,86	0,54	0,86
6	27,8	0,977	24,7	0,0765	3,26	0,805	0,58	0,857
7 (1)	13,3	0,657	30,4	0,13	2,44	0,774	0,705	0,782
7 (3)	18,7	0,43	27,3	0,984	3,43	0,784	0,644	0,791
7 (5)	24,6	0,904	24,8	0,158	4,52	0,801	0,617	0,802
7 (10)	36,9	1,09	24,6	0,119	6,75	0,824	0,56	0,837

Table B.17. Results of scenario 17

17_High Traffic_Negative Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	130	22,9	48,8	0,233	23,7	0,778	0,771	0,951
2	9,22	0,837	64,1	0,384	1,7	0,761	0,78	0,751
3	30	2,9	32,3	0,252	5,51	0,867	0,849	0,909
4	14,3	0,59	48,1	0,0857	2,63	0,845	0,753	0,85
5	77,80	8,81	38,80	0,03	14,30	0,85	0,85	0,85
6	37,50	3,88	32,50	0,33	6,92	0,93	0,88	0,73
7 (1)	27,2	1,85	33,1	0,181	4,99	0,853	0,841	0,856
7 (3)	36,1	2,59	29,8	0,16	6,61	0,848	0,849	0,856
7 (5)	42,9	1,28	28,8	0,142	7,87	0,847	0,852	0,856
7 (10)	56,9	3,67	28,3	0,118	10,4	0,846	0,847	0,85

Table B.18. Results of scenario 18

18_High Traffic_Positive Correlation_High Dispersion								
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	130	22,9	49,9	0,191	23,7	0,778	0,771	0,951
2	9,22	0,837	33,9	0,267	1,7	0,761	0,78	0,751
3	8,33	0,591	31,3	0,186	1,53	0,754	0,764	0,754
4	8,13	0,611	31,7	0,322	1,49	0,744	0,767	0,748
5	29,9	1,4	10	0	5,47	0,754	0,79	0,398
6	29,9	1,4	10	0	5,47	0,754	0,79	0,398
7 (1)	9,95	0,67	20,9	0,173	1,82	0,713	0,744	0,662
7 (3)	14	0,798	14,3	0,27	2,55	0,72	0,743	0,537
7 (5)	18,1	0,95	12,1	0,235	3,31	0,731	763	0,473
7 (10)	25,1	1,53	10,5	0,114	4,59	0,746	0,785	0,425

APPENDIX C: SCENARIOS WITH DIFFERENT DISTRIBUTION VARIANCES

Table C.1. Results of exponential arrival / lognormal (low variance) service time scenario

	Exp Arrival / LogN (Low Var.) Service							
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	40,9	8,3	62,8	0,0659	6,61	0,669	0,67	0,867
2	3,95	0,382	62,1	0,184	0,643	0,676	0,687	0,645
3	6,2	0,294	61,2	0,0497	1	0,76	0,652	0,723
4	4,48	0,455	61,3	0,0973	0,727	0,72	0,694	0,647
5	42	4,6	57,6	0,0843	6,83	0,86	0,544	0,753
6	19	1,41	57,7	0,0442	3,09	0,769	0,656	0,754
7 (1)	9,23	0,434	58,7	0,108	1,48	0,735	0,657	0,724
7 (3)	16,5	0,438	58	0,111	2,67	0,78	0,614	0,751
7 (5)	20,8	1,43	57,8	0,106	3,35	0,797	0,584	0,743
7 (10)	29,4	1,8	57,6	0,079	4,75	0,838	0,559	0,746

Table C.2. Results of lognormal (low variance) arrival / exponential service time scenario

	LogN (Low Var.) Arrival / Exp Service							
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	57,4	5,28	62,7	0,082	8,11	0,664	0,67	0,847
2	5,24	0,511	62	0,107	0,846	0,662	0,676	0,637
3	9,35	0,667	61,2	0,0332	1,51	0,767	0,652	0,723
4	6,8	0,764	61,2	0,0779	0,985	0,722	0,694	0,635
5	58	6,46	57,6	0,0611	9,38	0,849	0,546	0,747
6	28,9	2,68	57,7	0,045	4,67	0,766	0,653	0,75
7 (1)	12,6	0,815	58,9	0,0986	2,04	0,731	0,669	0,728
7 (3)	19,8	1,17	58,2	0,0964	3,21	0,767	0,635	0,743
7 (5)	26,1	1,56	57,9	0,0383	4,21	0,789	0,607	0,734
7 (10)	39,4	3,2	57,6	0,099	6,37	0,836	0,569	0,739

Table C.3. Results of lognormal (low variance) arrival / lognormal (low variance) service time scenario

	LogN (Low Var.) Arrival /LogN (Low Var.) Service							
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	41,1	4,58	62,8	0,0835	6,65	0,67	0,681	0,857
2	4,01	0,436	62	0,119	0,653	0,67	0,685	0,649
3	7,03	0,689	61,2	0,126	1,14	0,766	0,653	0,727
4	4,61	0,431	61,2	0,181	0,745	0,721	0,693	0,634
5	41,1	3	57,6	0,0814	6,68	0,857	0,551	0,746
6	23	3,57	57,7	0,0774	3,75	0,778	0,653	0,767
7 (1)	9,84	0,607	58,7	0,102	1,59	0,734	0,664	0,725
7 (3)	17	0,434	58,1	0,129	2,73	0,771	0,613	0,737
7 (5)	23,6	0,844	57,8	0,146	3,85	0,814	0,602	0,706
7 (10)	32,5	1,72	57,7	0,161	5,24	0,823	0,561	0,751

Table C.4. Results of exponential arrival / lognormal (high variance) service time scenario

	Exp Arrival / LogN (High Var.) Service							
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	75,5	22,1	62,8	0,0608	12,3	0,674	0,677	0,864
2	6,27	0,974	62,2	0,049	1,02	0,676	0,688	0,642
3	10,4	0,573	61,2	0,0846	1,69	0,766	0,661	0,729
4	6,39	0,364	61,3	0,125	1,03	0,718	0,694	0,639
5	64,50	5,60	57,60	0,05	10,50	0,86	0,56	0,76
6	34,20	4,68	57,70	0,06	5,57	0,78	0,66	0,76
7 (1)	14,8	2,4	58,9	0,0407	2,39	0,735	0,677	0,733
7 (3)	20,5	1,49	58,2	0,107	3,29	0,763	0,625	0,738
7 (5)	27,1	3,19	58	0,0861	4,39	0,779	0,607	0,75
7 (10)	39,6	3,61	57,8	0,0667	6,44	0,822	0,578	0,758

Table C.5. Results of lognormal (high variance) / exponential service time scenario

	LogN (High Var.) Arrival / Exp Service							
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	57,4	5,28	62,7	0,082	8,11	0,664	0,67	0,847
2	5,24	0,511	62	0,107	0,846	0,662	0,676	0,637
3	9,35	0,667	61,2	0,0332	1,51	0,767	0,652	0,723
4	6,08	0,764	61,2	0,0779	0,985	0,722	0,694	0,635
5	58,00	6,46	57,60	0,06	9,38	0,85	0,55	0,75
6	28,9	2,68	57,7	0,045	4,67	0,766	0,653	0,75
7 (1)	12,6	0,815	58,9	0,0986	2,04	0,731	0,666	0,728
7 (3)	19,8	1,17	58,2	0,0964	3,21	0,767	0,635	0,743
7 (5)	26,1	1,56	57,9	0,0383	4,21	0,789	0,607	0,734
7 (10)	39,4	3,2	57,6	0,099	6,37	0,836	0,569	0,734

Table C.6. Results of lognormal (high variance) / lognormal (high variance) service time scenario

	LogN (High Var.) Arrival /LogN (High Var.) Service							
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	65,9	6,71	62,8	0,0381	10,7	0,672	0,675	0,858
2	6,21	0,624	62,1	0,0447	1,1	0,67	0,658	0,644
3	10,9	1,52	61,2	0,0664	1,78	0,772	0,662	0,735
4	7,64	1,22	61,2	0,109	1,24	0,725	0,694	0,64
5	58,70	34,90	57,60	0,09	14,30	0,85	0,54	0,76
6	33,9	6,42	57,7	0,0929	5,48	0,772	0,645	0,73
7 (1)	13,3	1,36	58,9	0,15	2,15	0,739	0,673	0,725
7 (3)	22,5	3,45	58,3	0,0964	3,66	0,762	0,645	0,736
7 (5)	28,1	1,52	58	0,0708	4,53	0,787	0,609	0,732
7 (10)	39,7	2,3	57,8	0,106	6,39	0,813	0,569	0,74

Table C.7. Results of exponential / expoanential service time scenario

	Exp Arrival / Exp Service (Scenario 7)							
Policy No	Waiting Time	Half Width (0.95)	Av. Penalty	Half Width (0.95)	Avr Num in Queues	Util 1	Util 2	Util 3
1	49,3	1,49	62,7	0,0806	7,96	0,673	0,673	0,854
2	5,34	0,534	62,1	0,0731	0,868	0,675	0,685	0,65
3	8,57	1,11	61,2	0,0725	1,39	0,762	0,653	0,73
4	5,67	0,572	61,3	0,0752	0,916	0,722	0,691	0,644
5	50,30	5,32	57,60	0,03	8,13	0,85	0,54	0,75
6	25,8	2,35	57,7	0,048	4,18	0,77	0,656	0,747
7 (1)	11,9	2,7	59	0,857	1,92	0,731	0,667	0,731
7 (3)	19,4	1,85	58,2	0,0984	3,13	0,77	0,631	0,734
7 (5)	24,7	0,735	57,9	0,113	3,98	0,786	0,6	0,743
7 (10)	37,1	4,26	57,7	0,0778	6,2	0,833	0,565	0,77