AN OPTIMIZATION STUDY FOR URBAN PUBLIC TRANSPORT SYSTEM

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of
Dokuz Eylül University
In Partial Fulfillment of the Requirements for
the Degree of Master of Science in Industrial Engineering, Industrial Engineering Program

by
Özgür YALÇINKAYA

July, 2004
IZMIR

M.Sc THESIS EXAMINATION RESULT FORM

We certify that we have read this thesis and **"An Optimization Study for Urban Public Transport System"** completed by **Özgür YALÇINKAYA** under supervision of **Prof. Dr. G. Miraç BAYHAN** and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. G. Miraç BAYHAN

Supervisor

Prof. Dr. Edip TEKER

(Committee Member)

Yrd. Doç. Dr. Çiçek ÖZES

(Committee Member)

Approved by the

Graduate School of Natural and Applied Sciences

Prof. Dr. Cahit HELVACI

Director

ACKNOWLEDGEMENTS

I would first like to thank my advisor Prof. Dr. G. Miraç Bayhan for her continuing support, guidance and encouragement. I wish to give special thanks to Prof. Dr. Mahmut Bayhan for his comments and suggestions.

I would like to thank the staff of Izmir Metro A.Ş., and especially Emre Oral, for their understanding and help, also I would like to acknowledge Assistant Prof. Dr. Mehmet Çakmakçı for his help in providing communication with Izmir Metro A.Ş. management.

Finally, I would like to thank my dad, mother, elder sister and my fiancée Esen Gün for their never-ending supports, encouragements, helps and trust in me.

Özgür YALÇINKAYA

ABSTRACT

Because of the increasing car traffic and the need to improve the environment of the cities public transport systems have been constructed. Urban public transport providers have to make attractive the use of transport system for passengers. Most of the research in urban public transport area has been focused on the evaluation of system's performance, traffic management, vehicle planning and scheduling. In this study, for an underground urban public transport system in Izmir City, the aim is to find the optimum headways, to minimize the average passenger time spent in the metro-line with the requirement of fifty percent fullness rate of the carriages. For solution, first simulation models are constructed for each problem (weekday morning problem (*WMP*), weekday afternoon problem (*WAP*), Saturday problem (*STP*), and Sunday problem (*SNP*)). The data obtained from the simulation models are used for fitting metamodels, and then desirability functions are used for finding the optimum solution. The main contribution of the current study is to solve the optimization problem confronted in the urban public transport system by using response surface methodology (*RSM)*, and to provide direction for public transport providers to find out optimum solutions using *RSM*.

Keywords: Urban public transport systems, Simulation, Design of experiments, Metamodelling, Response surface methodology, Desirability functions

ÖZET

Artan araç trafiği ve şehirlerin çevresini iyileştirmek için kamu taşımacılık sistemleri inşa edilmiştir. Kamu taşımacılığını sağlayanlar taşımacılık sisteminin kullanılmasını yolcular açısından çekici hale getirmelidir. Kamu taşımacılığı alanında yapılan çalışmaların birçoğu sistem performansının değerlendirilmesi, trafik yönetimi, araç planlama ve çizelgeleme konularına odaklanmıştır. Bu çalışmada, İzmir Şehri metro kamu taşımacılık sistemi için, vagon doluluk oranının yüzde elli olması gerekliliği ile sistemde geçirilen ortalama yolcu süresini minimize edecek optimum tren sıklık sürelerinin bulunması amaçlanmıştır. Çözüm için, öncelikle her bir problem için simülasyon modeli oluşturulmuştur (haftaiçi sabah problemi (*HSP*), haftaiçi öğleden sonra problemi (*HÖP*), Cumartesi problemi (*CTP*), Pazar problemi (*PZP*)). Simülasyon modellerinden elde edilen veriler ile metamodeller oluşturulduktan sonra optimum çözüm için istek fonksiyonları kullanılmıştır. Bu çalışmanın ana katkısı, sehiriçi kamu taşımacılık sisteminde karşılaşılan optimizasyon probleminin çözümü için yanıt yüzey metodolojisinin (*YYM*) kullanılması, ve kamu taşımcılık hizmeti sağlayanlara *YYM*'ni kullanarak optimum sonuçları bulmaları için yol gösterici olmasıdır.

Anahtar sözcükler: Şehiriçi kamu taşımacılık sistemleri, Simülasyon, Deney tasarımı, Metamodelleme, Yanıt yüzey metodolojisi, İstek fonksiyonları

CONTENTS

Chapter Three

PROBLEM DEFINITION

Chapter Four

SIMULATION STUDY

Chapter Five
RSM STUDY

Chapter Six

CONCLUSION

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE
INTRODUCTION

1.1 Introduction

Because of the increasing car traffic and the need to improve the environment of the cities public transport systems have been constructed. Urban public transport providers have to make attractive the use of transport system for passengers. Most of the research in public transport area has been focused on the evaluation of system's performance, traffic management, vehicle planning and scheduling. Numerous approaches in the literature to tackle the problems of urban public transport have been proposed. Some of these approaches are algorithmic and can find the global optimum with respect to the goal function chosen, while the rest are heuristic and can find good solutions fast.

The literature can be classified by the tools which were used for reaching to the objective. Chiang et al.(1998) developed a rule-based expert system, Mackett & Edwards (1998) prepared questionnaire, Li (2000) build a simulation model, Priemus & Konings (2001) examined the other systems, Durmisevic & Sariyildiz (2001) used neural network method, and Carey & Carville (2003) developed a heuristic approach in their studies.

Chiang et al. (1998) described a knowledge-based railway scheduling system (*RSS*) for Taiwan Railway Administration's (*TRA*) railway scheduling operations. Scheduling process was divided into two levels, global scheduling and local scheduling. Railway scheduling problem was explained as determining the timetable that optimize a given criteria while satisfying the physical constraints with a given master scheduling plan (*MSP*) that included a set of train trips for serving passengers plan and railway facilities. Physical constraints were classified as running time, minimum stopover time, single-track-head on, overtaking, minimum headway, level crossing, and track assignment constrains. The scheduling objectives were in conflict, optimizing the passenger's service while minimizing the operation cost. *RSS* contained six models as data manager, automatic scheduler, database and knowledge base, user interface, schedule editor and performance evaluation modules. The global scheduler first generates an initial schedule for all train trips of current class according to the *MSP* without considering conflicts. The conflict finder then finds all of the conflicts and feeds the conflicts one by one to the local scheduler in time order (one of the three basic methods to arrange the order) for conflict resolution. Local scheduler can be viewed as a rule-based expert system. The rules that were used for resolving the conflicts detected by conflict finder were extracted from domain experts. Also, it was estimated that the current model saved %60 of the time and resources used for timetable preparation, and *RSS* can generate satisfactory, although not optimal, results in a short time.

Mackett & Edwards (1998) concerned with the way in which decisions are made about urban public transport systems, in particular the rationale underlying the decision-making process and the implications for the city in terms of travel demand, urban development and the environment. Their analysis was based upon a worldwide survey carried out as a part of a project to investigate the decision-making process involving the selection of the most appropriate technology for an urban transport system.

Li (2000) built a simulation model of a train station for passenger flow study. Simulation model included the processes, equipments and queues that, a passenger encounters from entering the station to exiting the station. All these encounters affected the total passenger travel times. Passenger processing time at various check points, queue time and queue length at the bottleneck areas, the number of people missing their first available train due to delays in the queue line, equipment utilization rate, and elapsed time between the first and the last passenger passed through a transfer point were used as evaluation criteria for determining an optimal solution. The aim was minimizing the total passenger travel times and increasing service quality.

Priemus & Konings (2001) denoted the needed conditions and strategies for introducing light rail system in the Netherlands, and focused on the opportunities for creating synergy between public transport and urban revitalization. They examined the systems in other developed countries such as France, Germany and Japan, and tried to learn from these systems. They added that developing a successful system depends on some factors as; an integrated approach (investment in public transport should be a part of urban policy), an associated policy (parking policy, road pricing and planning of park-and-ride areas), system structure of the public transport network, and last the quality of the urban and urban regional public transport.

Durmisevic & Sariyildiz (2001) realized a need for more systematic approach to the design and assessment of quality of underground spaces, because underground space has increasingly become a significant public domain for densely built urban areas, so that a better quality can be obtained. A detailed conceptual framework derived from the interviews that were fulfilled with different specialists. The aspects that determined the overall quality of underground spaces was grouped in three parts as; functional (accessibility, air quality, light, and temperature), psychological (safety, comfort, way finding, and attractiveness) and structural (dimensions, construction and separation walls, signing system) aspects. Two types of data were necessary, first type were related with the spatial characteristics that were actual and quantifiable values. Second type of data were obtained from questionnaires, analysis of data was done using the neural network method. They proposed an approach for a consistent assessment of these factors so that in future it can be integrated into a decision support system that can help indicate problem in existing underground spaces and offer support to architects designing new underground spaces.

Carey & Carville (2003) focused on busy complex rail stations, which were key components of the busy passenger rail networks, and were the location of most train conflicts. Train planning for a large busy station included drawing up a schedule to ensure that there are no conflicts between any trains, while ensuring that all the requirements and constraints are satisfied and minimizing any deviations or cost of deviations from desired or preferred times, platforms or lines for each train. An algorithmic approach was developed for generating a station schedule, which is a scheduled arrival time, scheduled departure time and platform allocation for each train. The developed scheduling algorithms can be used for train scheduling for a rail line or network.

Sinclair & Oudheusden (1997), Wu & Hounsell (1998), Ferrari (1999), Şahin (1999), Parkes & Ungar (2001), Huisman et al. (2002) developed mathematical models as a tool.

Sinclair & Oudheusden (1997) proposed a minimum cost network flow model to deal with the problem of bus trip scheduling in heavily congested cities. They called the mechanism, which chosen for representing the bus departures from the terminals, as trip frequency scheduling. Trip frequency scheduling specifies the number of buses per time period that must leave the terminal during different time intervals in the day. They constructed a network optimization model for the trip frequency-scheduling problem and used goal programming to solve the problem. Two objectives were determined, schedule a number of buses as close as possible to the number of trips required in each time interval (a measure of the quality of service) and minimize the total travel time (a measure of the operating cost). Constraints were, parking space at the terminals, number of buses available at a depot, minimum service levels and bus capacity. Although in the classic scheduling problem bus demand must be met unconditionally, in their approach a bus can be cancelled if its cost is too high.

Wu & Hounsell (1998) illustrated an analytical approach for the pre-implementation evaluation of pre-signals. Pre-signals concept was a bus priority strategy which aims to give buses priority access into a bus advanced area of the main junction stop line for avoiding the traffic queue and also for reducing bus delay at the signal controlled junction. Traffic signals were installed at or near the end of a with-flow bus lane to provide buses with priority access to the downstream junction. Analytical equations were developed for determining optimum pre-signal timings, signal time settings, delay savings, lengths of the relocated traffic queue, and required bus advance area for designing pre-signalised junctions and estimations of delay.

Ferrari (1999) dealt with a bimodal transport system and proposed a new method for solving the programming problem with equilibrium constraints. Decision variables were; road pricing, transit ticket price, and service characteristics of transit, also the constraints were, physical and environmental capacity constraints, and budget constraints. The aim was choosing the values of decision variables, which maximize the average user satisfaction with the satisfied budget and capacity constraints and the system is in equilibrium.

Şahin (1999) dealt with inter-train conflicts problem. An inter-train conflict (meet/pass) simply occurs when two opposing trains move on a single-track section between neighbouring meet points, or if a faster train catches a slower one moving in the same direction. Şahin dealt with analyzing dispatchers' decision process in inter-train conflict resolutions and developing a heuristic algorithm for rescheduling trains by modifying existing meet/pass plans in conflicting situations in a single-track railway. A zero-one mixed-integer-programming model was built, the mathematical model of rescheduling process in railway traffic control problem is similar to job-shop scheduling problem. While the objective was to minimize the sum of job completion times in the job-shop scheduling, the sum of running times of trains were minimized in the rescheduling process. Also, a heuristic algorithm was developed in order to obtain better conflict solutions than train dispatchers and optimal or near optimal solutions in reasonable length of time. Then, three solution methods; optimal solution (was found by the software LINDO), dispatcher's solution and heuristic's solution were compared. The comparison criteria were; measure of effectiveness, total waiting times, computation time. The heuristic gave better solution than dispatcher's, also the heuristic algorithm performs almost as well as the optimal solution method in selecting the better conflicting train to stop.

Parkes & Ungar (2001) presented a computational study of an auction-based method for decentralized train scheduling, which was well suited to the natural information and control structure of modern railroads. They assumed separate network territories, with an autonomous dispatch agent responsible for flowing of the trains over each territory. Each train was represented by a self-interested agent, which bids for the right to travel across the network from its source to destination, submitting bids to multiple dispatch agents along its route as necessary. The used bidding language allowed trains to bid for the right to enter and exit territories at particular times, and also to represent indifference over a range of times. The problem was to compute a robust and safe meet/pass schedule for the movement of trains over the network to maximize the total cost-adjusted value over all trains. The global objective was to find a safe schedule that maximizes the total net value (total value minus cost of delay across all trains that run). In addition, the input was a set of trains, each with a defined route over a track network, a value for completing its journey, and an optimal departure and arrival time and cost function for off-schedule performance. They formulated the winner determination problem as a mixed-integer program. Computational results on a simple network with straightforward best-response bidding strategies demonstrated that the auction computed near-optimal system-wide schedules.

Huisman et al. (2002) developed an analytical tool for railway networks that not required timetables, only the inputs were train frequencies. Due to absent detailed train schedules of new railway networks, the performance of them could not be measured or simulated. The main objective was to develop an analytically tractable queuing network model for total railway networks, taking into account dependencies and interaction between the individual components that are stations, junctions and section tracks.

In addition, some other tools like fuzzy multi-criteria analysis, job shop scheduling formulation, a local feedback-based travel advanced strategy are used.

Kreuger et al. (1997) described a novel constraint model for scheduling train trips on a network of tracks used in both directions. They thought scheduling train trips on a network of tracks is an optimization problem that resembles but also differs in certain ways from other typical scheduling tasks, therefore this kind of problem can be modelled as a job-shop scheduling problem. Train trips assumed as jobs to be scheduled on tracks, tracks as resources, and each train trip traversing a track represents a task, and also the traversal time was taken as the duration of the task. The problem is stated as, scheduling a set of train trips over a fixed network of predetermined paths where trains travel in both directions on single tracks, connecting nodes where trains can meet and overtake and maintaining reasonable bounds on waiting and total times. They also noted that although the used simple solver in the performance tests does not lend itself well to optimization, the founded schedules are quite close to manually found optima in general.

Yeh et al. (2000) presented an effective fuzzy multi-criteria analysis approach to performance evaluation for urban public transport systems that involving multiple criteria of multilevel hierarchies and subjective assessments of decision alternatives. Approach provided a structured framework for the decision makers to think and handle the performance evaluation problem systematically from a wide variety of viewpoints in addition to the traditional safety and operation criteria. They applied this approach to a case study of an urban public transport system operated by ten bus companies, in Taiwan. Five criteria were selected for evaluating the performance of ten bus companies. The criteria were, safety (sub-criteria were accident rate, average vehicle age, average vehicle breakdown and traffic offence rate), comfort (air-conditioned vehicle rate, passenger information, vehicle cleanliness, seat comfort, driver's driving skills, driver's appearance and driver's friendliness were selected as sub-criteria), convenience (the selected sub-criteria were punctuality of the bus service, route transferability, terminal space and service reliability), and operation social duty (cost effectiveness and service efficiency were sub-criteria).

Dorfman & Medanic (2003) developed a local feedback-based travel advanced strategy (*TAS*) by using a discrete event model of train advances along lines of the railway. *TAS,* which was used for scheduling trains in a railway network, was a service discipline at each meet and pass node determining, which of the trains in the vicinity should continue to travel, and which should be stopped at the meet and pass node. *TAS* has some computational advantages as computing more information then other approaches, can be used to quickly develop schedules for perturbed cases (as change in a particular departure time, existence of a behind the schedule train), also operating way is similar to train dispatchers approach. The three performance criteria that related to measuring the performance of a *TAS* were; time to clear the line (the efficiency ratio), delay of all trains, and the maximal delay. Time to clear the line was particularly appropriate to assess the degree of optimality of a schedule when the *TAS* was employed. In addition, generally *TAS* does not generate mathematically optimal solutions, *TAS* develops suboptimal schedules that closely approach the optimal in practical situations. Also, *TAS* can be used to develop an energy-efficient schedule for a railway network.

In addition, a review paper Cordeau et al. (1998) classified the optimization models for transportation problems. The problems were categorized in two main classes, routing problems and scheduling problems.

The briefly mentioned papers are shown in Table 1.1 with the used tools, problem type and the objective of studies.

Because of its public structure, the major goal of the Izmir Metro company is to give satisfactory, cheap and fast service to passengers. Thus, minimizing the time that a passenger spends in the line is very important.

In this study, the problem is to find the headways (input factors), for each ten-time interval in five days between Monday and Friday, for each four-time interval in Saturday, and for each three-time interval for Sunday.

Table 1.1 Objectives, problem types and the used tools in literature

Paper	Tool	Objective	Problem type
Sinclair & Oudheusden (1997)	mathematical prog. (goal prog.)	to schedule bus trips and to minimize the total travel time	minimum cost network flow model
Kreuger et al. (1997)	job shop scheduling formulation	to schedule train trips	novel constraint model on a network of tracks
Mackett & Edwards (1998)	questionnaire	to prepare a decision support system	evaluation of expectations
Chiang et al. (1998)	a rule-based expert system	to optimize the service of passenger and to minimize the operation cost	knowledge-based railway scheduling
Wu & Hounsell (1998)	analytical model	to determine optimum pre-signal timings, signal time settings, delay savings, lengths of the relocated traffic queue.	pre-implementation evaluation of pre-signals
Şahin (1999)	mathematical prog. (mixed-integer prog.)	to analyze dispatchers' decision process in inter-train conflict resolutions and to develope a heuristic algorithm for rescheduling trains	inter-train conflict resolutions
Ferrari (1999)	mathematical prog.	to maximize the average user satisfaction with the satisfied budget and capacity constraints	programming problem with equilibrium constraints
Li (2000)	simulation	to minimize the total passenger travel times and to increase service quality	passenger flow study
Yeh et al. (2000)	fuzzy multi-criteria analysis	to evaluate the performance of bus companies	performance evaluation of urban public transport
Durmisevic & Sariyildiz (2001)	neural networks	to prepare a decision support system	design and assessment of quality of underground spaces
Priemus & Konings (2001)	examination of the other systems	to focus on the opportunities for creating synergy between public transport and urban revitalization	introducing light rail system
Parkes & Ungar (2001)	mathematical prog. (mixed-integer prog.)	to find a safe schedule that maximizes the total net value	auction-based method for decentralized train scheduling
Huisman et al. (2002)	analytical model	to develop an analytically tractable queuing network model for total railway networks for performance evaluation	performance evaluation of railway networks
Carey & Carville (2003)	heuristic approach	to generate a station schedule	train conflicts
Dorfman & Medanic (2003)	a local feedback-based travel advanced strategy	to schedule trains	a local feedback-based travel advanced strategy

The objective is to minimize the average passenger time spent in the metro-line (the first response) with the requirement that the fullness rate of the carriages as fifty percent (the second response). We solve this problem by integrating Response Surface Methodology (*RSM*) into simulation. To the best of our knowledge, this is the first study that uses *RSM* and simulation integration for solving a urban public transport problem.

The thesis is organized as follows. In chapter two, after explaining the response surface methodology a related literature review is given. Problem in Izmir Metro is defined in detail and the system is introduced in chapter three. In chapter four, working logics of the system and simulation model are explained. Also, input data for model, assumptions before coding phase of the simulation model, flowcharts for occurred events, some important attributes and variables that are used in the simulation model are given in chapter four. In addition, a flowchart that denotes modelling logic of the Halkapinar station is demonstrated. At the end of the chapter four, the verification and validation techniques that are used for simulation model are denoted. In chapter five, first *RSM* study is explained, then the optimization points are searched for four problems, which are weekday morning (*WMP*), weekday afternoon (*WAP*), Saturday (*STP*) and Sunday (*SNP*) problems. After developing validated metamodels Derringer-Suich multi-response optimization procedure is used for these four problems.

CHAPTER TWO
RESPONSE SURFACE METHODOLOGY

2.1 Introduction to Response Surface Methodology

Response surface methodology (*RSM*) is a collection of mathematical and statistical techniques, which are useful for developing, improving, and optimizing processes. It also has important applications in the design, development, and formulation of new products, as well as in the improvement of existing product designs, and it is an effective tool for constructing optimization models (Myers & Montgomery, 1995, p.1).

RSM consists of the experimental strategy for exploring the space of the process or input factors, empirical statistical modelling to develop an appropriate approximating relationship between the yield and the process variables, and optimization methods for finding the levels or values of the process variables that produce desirable values of the response outputs (Myers & Montgomery, 1995, p.3).

RSM was proposed by Box and Wilson in 1951 for finding the input combination that minimizes the output of a real, non-simulated system. Then, it was generally used for random simulation models (Angün et al., 2002, p.377).

Box and Wilson (1951) laid the foundations for *RSM*. That paper was important not only because it described what became an entire field of research for the next fifty years but also because it changed dramatically the way that engineers, scientists, and statisticians approached industrial experimentation. They outlined a sequential philosophy of experimentation that encompasses experiments for screening, region seeking (such as steepest ascent), process/product characterization, and process/product optimization. Clearly, *RSM* includes much more than second order model fitting and analysis. Indeed, *RSM*, broadly understood, has become the core of industrial experimentation (Myers et al., 2004, p.53).

RSM is generally used for optimizing the performance (or a model) of an unknown system, which is subject to controllable, uncontrollable and unknown variables. *RSM* can be applied to any system that has key elements like a criterion of effectiveness known as the response of the system, which is measurable on continuous scale, and quantifiable independent variables that affect the system's performance. At these conditions, *RSM* is a group of techniques for finding the optimum response of the system in an optimum fashion. Although from a historical viewpoint *RSM* has been applied primarily to Operations Management, its greatest value to decision scientists is its potential application to simulation studies (Brightman, 1978, p.481).

There are two distinct phases in *RSM*. If the values of the decision variables are thought to produce a level of the system performance far below the maximum (minimum) , the method of steepest ascent (descent) is employed to reach the optimum region quickly. The second phase employs canonical analysis to determine the exact values for decision variables that ensure optimum system performance (Brightman, 1978, p.482).

Graphical view of a response surface and contour plot for a maxima problem with two independent variables is shown in Figure 2.1. In addition, current operating condition and maximum response point are denoted on contour plot.

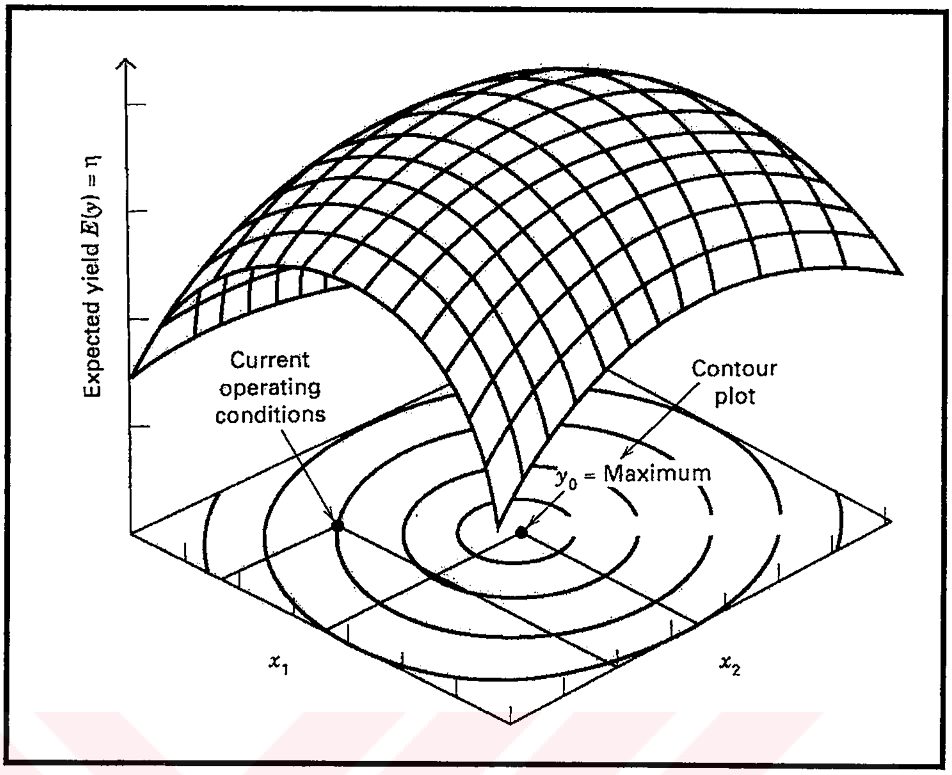

Figure 2.1 A three dimensional response surface

2.2 Approximating Response Functions

For example, if the objective is to find the levels of independent variables ($\phi_1, \phi_2, ..., \phi_k$) that maximize the response (y) of a specific process, the process should be identified as a function of independent variables, which is;

$$y = f(\phi_1, \phi_2, ..., \phi_k) + \varepsilon \tag{2.1}$$

where ε represents the noise or error that is observed in the response (y). Since the error term should be normally, identically and independently distributed with zero mean and constant variance (σ^2), the expected response is;

$$E(y) = E[f(\phi_1, \phi_2, ..., \phi_k)] + E(\varepsilon) = \eta \tag{2.2}$$

and then the surface represented by;

$$\eta = f(\phi_1, \phi_2, ..., \phi_k) \tag{2.3}$$

is called a response surface. The variables $\phi_1, \phi_2, \ldots, \phi_k$ in equation (2.3) are called natural variables, because they are expressed in the natural units of measurement. In much *RSM* work it is convenient to transform the natural variables to coded variables $x_1, x_2, \ldots, x_k$ where these coded variables are usually defined to be dimensionless with mean zero and the same spread or standard deviation, and also fall between –1 and +1. The formulation, which is used for transforming natural variables to coded variables is;

$$x_k = \frac{\phi_k - [\max(\phi_k) + \min(\phi_k)]/2}{[\max(\phi_k) - \min(\phi_k)]/2} \tag{2.4}$$

and the response surface in terms of coded variables is;

$$\eta = f(x_1, x_2, \ldots, x_k) \tag{2.5}$$

(Myers & Montgomery, 1995, p.3).

In most *RSM* problems, the form of the relationship between the independent variables and the response is unknown, it is approximated. Thus, the first step in *RSM* is to find an appropriate approximation for the true functional relationship between response and the set of independent variables. Usually, a low-order polynomial in some region of the independent variables is employed. If the response is well modelled by a linear function of the independent variables, then the approximating function is the first order model;

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \tag{2.6}$$

If there is curvature in the system, then a polynomial of higher degree must be used, such as the second order model;

$$y = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \sum_{j=1}^{k} \beta_{jj} x_j^{\,2} + \sum_{i}^{k-1} \sum_{j}^{k} \beta_{ij} x_i x_j + \varepsilon \tag{2.7}$$

where $i = 1,2,\ldots,k\text{-}1$ and $j = 1,2,\ldots,k$ also $i<j$.

Almost all *RSM* problems utilize one or both of the approximating polynomials mentioned above. However, it is not the case that a polynomial model will be a reasonable approximation of the true functional relationship over the entire space of the independent variables. These models usually work quite well only for a relatively small region. The method of least squares is used to estimate the parameters in the approximating polynomials.

The response surface analysis is then performed using the fitted surface. If the fitted surface is an adequate approximation of the true response function, then the analysis of fitted surface will be approximately equivalent to analysis of the actual system. This type of model is called as metamodel. The model parameters can be estimated most effectively if proper experimental designs are used to collect the data. Designs for fitting response surfaces are called response surface designs. The response surface design used for fitting response surface directly affects the quality of the estimate (Montgomery, 2001, pp.427-429).

The second order models are widely used in *RSM* because; it has a flexible structure that means it can take on a wide variety of functional forms, so it will often work well as an approximation to the true response surface, and estimating the regression parameters in the second order model is easy, also there is considerable practical experience indicating that second order models work well in solving real response surface problems (Myers & Montgomery, 1995, p.7).

2.2.1 The Metamodel Concept

Metamodelling is a process of developing a mathematical relationship between a response measure of interest and a set of input variables (Batmaz & Tunali, 2003, p.455).

Metamodel first described by Blanning (1975) as a mathematical relationship between one or more sets of sensitivity measures of interest and the sets of input to the metamodel (Aytuğ et al., 1996, p.24).

Let δ_j denote a factor j influencing the outputs of the real-word system (j = 1,2,…,s), and let W_c denote the system response vector (c = 1,2,…,w). Without loss of generality, the discussion can be simplified by considering a system with a single response W, since a multiple response system can be considered as a set of single response systems. The relationship between the response variable W and the inputs δ_j of the system is represented by;

$$W = f_1(\delta_1, \delta_2, \ldots, \delta_s) \tag{2.8}$$

A simulation model is then an abstraction of the real system, in which it is considered only a selected subset of the input variables $\{\delta_j \mid j = 1,2,\ldots,r\}$ where r is significantly smaller then the unknown s. the response of the simulation W' is then defined as a function f_2 of this subset and a vector of random numbers v representing the effect of the excluded inputs;

$$W' = f_2(\delta_1, \delta_2, \ldots, \delta_r, v) \tag{2.9}$$

A metamodel is a further abstraction, in which a subset of the simulation input variables are selected $\{\delta_j \mid j = 1,2,\ldots,m,\ m \leq r\}$ and the system described as;

$$W'' = f_3(\delta_1, \delta_2, \ldots, \delta_m) + \varepsilon \tag{2.10}$$

where ε denotes a fitting error, which has an expected value of zero. These levels of abstraction are shown in Figure 2.2 (Yu & Popplewell, 1994, p.788).

The simulation model is leaner than the real-word system, it contains fewer variables and these are all under the control of the experimenter. The simulation model, although simpler than the real-word system, is still a very complex way of relating input to output. A simpler analytical model may be used as an auxiliary to the simulation model in order to better understand the more complex model and to provide a framework for testing hypotheses about it. This auxiliary model is referred to as a metamodel (Friedman, 1995, pp.15-20).

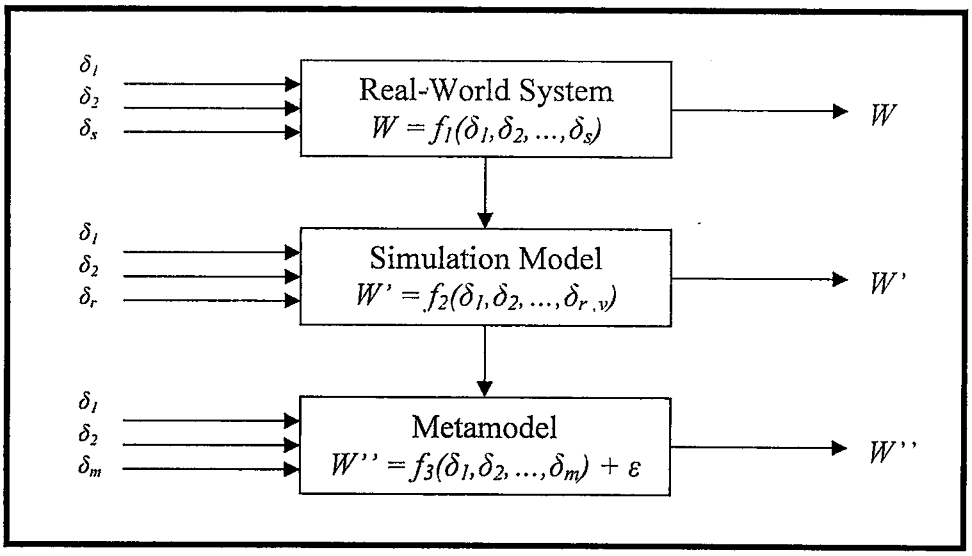

Figure 2.2 Metamodel concept

Metamodelling provides one approach to statistical summarization of simulation results, allowing some extrapolation from the simulated range of system conditions and therefore potentially offering some assistance in optimization. While simulation model is an abstraction of the real system, metamodel is an abstraction of simulation model; so metamodel is a further abstraction of the real system (Yu & Popplewell, 1994, p.778).

The objective of building a metamodel is to determine a (relatively simple) functional relationship between the system response and selected decision variables. Thus, it becomes much easier (cheaper) not only analyzing the simulation output, but also predicting how the real system will react to specify combinations of the set of controllable input variables. It is also straightforward to perform sensitivity analysis of the simulation model parameters and "what-if" questions without having to perform additional simulation runs (Santos & Nova, 1999, p.502).

The purpose of a metamodel is to estimate or approximate the response surface. A metamodel is often specified to be a regression model where the independent variables for the regression are the simulation input parameters and the dependent variable is the response of interest. Then the metamodel, instead of the actual simulation model, can be used economically to learn about how the response surface would behave over various regions of input-factor space and thus to estimate how the response would change at a particular point of the input factors are changed slightly or perhaps to find approximately optimal settings of the input factors (Sridharan & Babu, 1998, p.592).

Some of the benefits of the metamodel are;

- Simulation models are flexible and solve real problems without making too many restricting assumptions as in most analytical models. By using metamodels in post simulation analysis a good approximation to reality is provided.
- Through the use of experimental design the number of simulation runs required to generate a metamodel is drastically reduced. This also reduces the amount of time and effort that is spent in conducting simulation. Thus simulation becomes an attractive technique to use.
- The use of metamodels allows some generalizations to simulation output. However, these generalizations have to be within the defined boundaries of the problem. Sensitivity analysis on model parameters can also be easily carried out without re-running costly simulation programs.
- The model is valid and yields satisfactory solutions that are comparable to simulation results.
- Regression metamodels are usually simpler and easier to use than most analytical models (Madu, 1990, p.388).

2.3 The Followed Steps by a *RSM* Study

A typical *RSM* study consists of the following steps;

A) Estimation Process

a) The Studies Before Estimating the Metamodel

- Determining the objectives of developing a metamodel for the simulation model.

- Identifying all of the decision variables and their characteristics such as being discrete or continuous, and the performance measure(s).
- Identifying the operability regions of all the decision variables. The developed and validated simulation model that represents the behaviour of real system is used for this purpose.
- Identifying the accuracy of the metamodel with respect to the selected performance measure(s).
- Determining the validity measure(s) of the metamodel.
- Applying the factor screening. In this phase, the important decision variables (i.e., the decision variables which have statistically more effects on the performance measures) are determined.

b) The Studies for Estimating the Metamodel

- According to the factor screening results, build a 2^k full factorial or 2^{k-p} fractional factorial with centre points design for estimating the metamodel.
- Determining the low and the high levels for each decision variable. The operability region of the related decision variable should cover the low and high level of the related decision variable. The low and the high levels for the decision variables are coded as −1, +1 respectively.
- Making tactical decisions on executing the simulation model such as specifying the number of replications at each design point, the length of the replication, and the variance reduction technique which will be used.
- After executing the model at each design point, applying the least square estimation procedure for fitting the response to a first order model.
- Analyzing significance of the model parameters, main and interaction effects, lack-of-fit, and curvature effects. If the first order model is proper according to the results, the optimization process starts. If it is not proper, the first order model with interaction or second order model is fitted.

B) Optimization Process

- Determining the gradient vector of the performance measure, if the first order model is proper. This gradient vector is used in one of the process improvement techniques such as Steepest Ascent/Descent. In these methods, the centre of the design is changed to increase the performance of the system. If there is no improvement, the method stops for estimating a new first order model in new ranges, then, same procedure which was explained will be followed. This process continues until the new design has the interaction or curvature effects. If it is so, the interaction or second order model is fitted.

- After fitting second order model, applying the canonical analysis for calculating the optimal levels of the decision variables (Yildiz, 2003, pp.41-43).

2.4 The Sequential Nature of *RSM*

Most applications of response surface methodology are sequential in nature. That means, at first, some ideas are generated concerning which factors or variables are likely to be important in the response surface study. This usually leads to an experiment designed to investigate these factors with a view toward eliminating the unimportant ones. This type of experiment is usually called a screening experiment. The objective of factor screening is to reduce the list of candidate variables to a relatively few, so that subsequent experiments will be more efficient and require fewer runs or tests. A screening experiment is referred as *phase zero* of a response surface study.

After identifying important independent variables *phase one* of the response surface study begins. In phase one, the objective is to determine if the current levels or settings of the independent variables result in a value of the response that is near the optimum, or if the process is operating in some other region that is (possibly) remote from the optimum. If the current settings or levels of the independent variables are not consistent with optimum performance, then the experimenter must determine a set of adjustments to the process variables that will move the process toward the optimum. Phase one of response surface methodology makes considerable use of the first order model and an optimization technique called the method of steepest ascent/descent.

Phase two of a response surface study begins when the process is near the optimum. At this point the experimenter usually wants a model that will accurately approximate the true response function within a relatively small region around the optimum.

Because the true response surface usually exhibits curvature near the optimum, a second order model (or perhaps some higher order polynomial) will be used. Once an appropriate approximated model has been obtained, this model may be analyzed to determine the optimum conditions for the process (Myers & Montgomery, 1995, pp.10-11).

This sequential experimental process is usually performed within some region of the independent variable space called the operability region. This sequential nature is denoted in Figure 2.3.

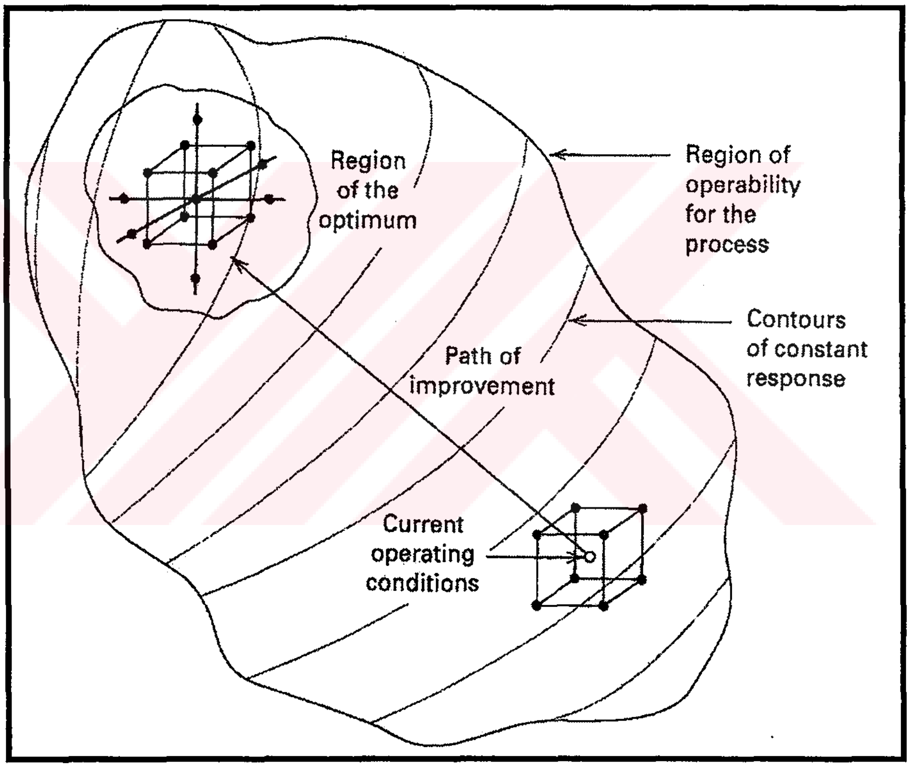

Figure 2.3 The sequential nature of *RSM*

2.5 The Method of Steepest Ascent/Descent

RSM is hill climbing (for a maxima objective), and the objective is locating the mountain summit. As in most mountain-climbing expeditions, the base camp is far below the summit. The objective of the method of steepest ascent is to move swithly up the mountain without stopping for exploring the lower peaks (Brightman, 1978, p.482).

Frequently, the initial estimate of the optimum operating conditions for the system will be far from the actual optimum. At this situation, the objective is to move rapidly to the general vicinity of the optimum. The method of steepest ascent is a procedure for moving sequentially along the path of steepest ascent, that is, in the direction of the maximum increase in the response. If minimization is desired, then this technique is called as the method of steepest descent method.

As an example, the region of fitted first order response surface and the path of steepest ascent for a response with two independent variables (x_1 and x_2) is shown in Figure 2.4. The fitted first order model is as;

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i \qquad (2.11)$$

and the contours of $\hat{y}$ is a series of parallel lines as denoted in Figure 2.4.

The direction of steepest ascent is the direction in which response increases most rapidly. This direction is parallel to the normal to the fitted first order response surface. The coordinates along the path of steepest ascent depend on the nature of the regression coefficients in the fitted first order model (Montgomery, 2001, p.430).

The method of steepest ascent contains the following steps;

- Fitting a first order model by using an orthogonal design. Two-level designs will be quite appropriate, although centre runs are recommended.

  Consider a situation in which N experimental runs are conducted on k design variables $x_1, x_2, \ldots, x_k$ and a single response y. A model is postulated of the types;

$$y = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i \,, \; (i = 1,2,\ldots,N) \qquad (2.12)$$

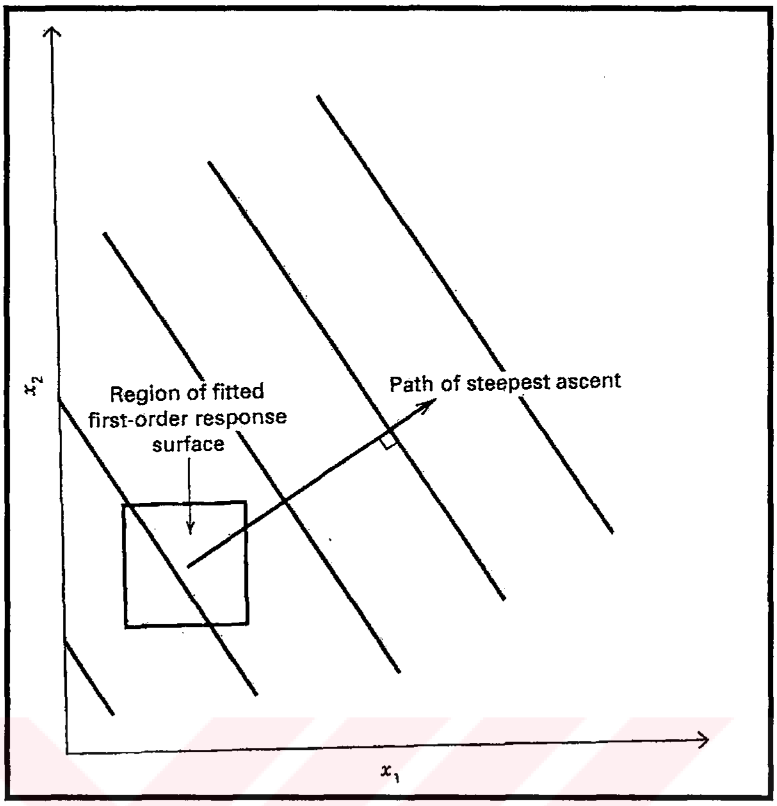

Figure 2.4 Path of steepest ascent

and the fitted model is given by;

$$y = b_0 + b_1 x_{i1} + b_2 x_{i2} + \ldots + b_k x_{ik} \tag{2.13}$$

where b_i are found by the method of least squares. For the first order model and a fixed sample N, if $x_j \in [-1,+1]$ for $j = 1,2,\ldots,k$, then $Var\left(b_i / \sigma^2\right)$ for i = 1,2,…,k is minimized if the design is orthogonal and all x_i levels in the design are ± 1 for i = 1,2,…,k (Myers & Montgomery, 1995, pp.283-284).

- Computing a path of steepest ascent (descent) if maximizing (minimizing) response is required.
- Conducting experimental runs along the path. That is, doing single or replicated runs and observing the response value. The results will normally show improving values of response. At some region along the path, the improvement will decline and eventually disappear. This stems from the deterioration of the

simple first-order model once one strays too far from the initial experimental region.

- Choosing a base at some location for a second experiment, where an approximation of the maximum (or minimum) response is located on the path. Again, the design should be a first order design. It is quite likely that centre runs for testing curvature, and degrees of freedom for interaction-type lack-of-fit are important at this point.
- Conducting a second experiment and fitting another first order model. Making a test for lack-of-fit. Computing a second path based on the new model, if lack-of-fit is not statistically significant. Conducting single or replicated experiments along this second path. It is quite likely that the improvement will not be as strong as that in the first path. After improvement is diminished, one has a base for conducting a more elaborate experiment and a more sophisticated process optimization (Myers & Montgomery, 1995, p.184).

A general algorithm for determining the coordinates of a point on the path of steepest ascent (with the assumption that the points x_i are the base or origin point, namely $x_1 = x_2 = ,\ldots, = x_k = 0$) is as follows;

- Choosing a step size in one of the variables, say Δx_j. Usually, the variable that is known most about or has the largest absolute regression coefficient $|\hat{\beta}_j|$ is selected.
- The step sizes in the other variables are;

$$\Delta x_i = \frac{\hat{\beta}_i}{\hat{\beta}_j / \Delta x_j} \qquad (2.14)$$

- Converting the Δx_i from coded variables to the natural variables (Montgomery, 2001, pp.435-436).

As the experimental region moves near the region of optimum conditions, it is certainly expected that curvature would be more prevalent. If a test for curvature finds significant quadratic terms, it would be suspected that the steepest ascent (descent) methodology would became effective. At this point, the investigators will surely be interested in finding optimum conditions through the use of a fitted second order model (Myers & Montgomery, 1995, pp.188-189).

2.6 Analysis of Second Order Response Surface

2.6.1 Central Composite Designs (*CCDs*)

When the experimenter is relatively close to the optimum, a model which includes curvature is usually required to approximate the response. In most cases, the second order model is adequate.

Second order models cannot be fitted with two-level designs plus centre points. The minimum conditions to fit a second order model are, (a) at least $1+2k+k(k-1)/2$ distinct design points, where k is the number of design variables, and (b) at least three level of each design variable. In the case of first order designs the dominant property is orthogonality. In the case of second order designs, orthogonality ceases to be such an important issue; and estimation of individual coefficients, while still important, becomes secondary to the scaled prediction variance ($N\,Var\,\hat{y}(x)/\sigma^2$) (Myers & Montgomery, 1995, p.297).

For fitting a second-order model there is a class of designs, central composite designs (*CCDs*) are common. *CCDs* are two-level full or fractional factorial designs that have been augmented with a small number of carefully chosen treatments to permit estimation of the second order response surface models (Neter et al., 1996, p.1281).

CCDs are obtained from resolution V, full or fractional factorial designs by the adding of star points and perhaps more centre points. Star points are points where one of the x_i takes on the values $\pm\,\alpha$ while the remaining x_i are all zero (Hood & Welch, 1993, p.117).

Resolution V designs are designs in which no main effect or two-factor interaction is aliased with any other main effect or two-factor interaction, but two-factor interactions are aliased with three-factor interactions (Myers & Montgomery, 1995, p.139).

The total number of experimental trials planned, that is indicated by n_T is;

$$n_T = 2^{k-f} n_c + 2^k n_s + n_o \qquad (2.15)$$

and the α that generates a rotatable design is given by;

$$\alpha = \left[\frac{2^{k-f}(n_c)}{n_s} \right]^{1/4} \qquad (2.16)$$

where k is the number of factors, f is the level of fractional in the two-level factorial design selected, n_c is the number of replications at each design point, n_s is the number of replications at each star points, n_o is the number of replications at the centre point, and α is the axial distance.

A rotatable design is one for which $N\ Var\ \hat{y}(x)/\sigma^2$ has the same value at any two locations that are the same distance from the design centre. In other words, $N\ Var\ \hat{y}(x)/\sigma^2$ is constant on spheres (Myers & Montgomery, 1995, p.306).

While rotatability is a desirable property of a *CCD*, it should not be the sole basis for making the choice of α, in some situations it may be physically impossible or difficult to extent the star points beyond the experimental region defined by the upper and lower limits of each factor, where α must not exceed 1 ($\alpha = 1$ is often called a face-centered design). In this circumstance the resulting lack of rotatability may not be considered a serious disadvantage (Neter et al., 1996, pp.1283-1287).

The three components of *CCD* play important roles in building second order models:

- The resolution V fraction contributes in a major way in estimation of linear terms and two factor interactions. It is variance-optimal for these terms. The factorial points are the only points that contribute to the estimation of the interaction terms.
- The axial points contribute in a large way to estimation of quadratic terms, and do not contribute to the estimation of interaction terms.
- The centre runs provide an internal estimate of error (pure error) and contribute toward the estimation of quadratic terms (Myers & Montgomery, 1995, p.298).

The natural competitor for the *CCD* is three-level (3^k) factorial design. Actually, 3^2 design is a face centre cube, thus a *CCD*. But when k becomes large, the 3^k factorial design includes an excessive number of design points. For $k>3$ the number of design points for the 3^k design is usually considered impracticable for most applications, due to that *CCD* is the most popular class of second order designs (Myers & Montgomery, 1995, p.318).

2.6.2 Location of the Stationary Point

The aim is to find the levels of $x_1, x_2, \ldots x_k$ that optimize the predicted response. This point, if it exists will be the set of $x_1, x_2, \ldots, x_k$ for which partial derivates are zero, namely;

$$\partial\hat{y}/\partial x_1 = \partial\hat{y}/\partial x_2 = \ldots = \partial\hat{y}/\partial x_k = 0 \quad (2.17)$$

This point, say $x_{1,s}, x_{2,s}, \ldots, x_{k,s}$ is called the stationary point (Montgomery, 2001, p.439).

The fitted second order model in matrix notation is;

$$\hat{y} = b_0 + x'b + x'\hat{B}x \quad (2.18)$$

where b_0 contains estimate of the intercept, b contains estimate of the linear, and $\hat{B}$ contains estimate of the second order coefficients. Actually;

$$x' = [x_1, x_2, \ldots, x_k] \quad (2.19)$$

$$b' = [b_1, b_2, \ldots, b_k] \quad (2.20)$$

$$\hat{B} = \begin{bmatrix} b_{11} & b_{12}/2 & \ldots & b_{1k}/2 \\ & b_{22} & \ldots & b_{2k}/2 \\ & & \ldots & \ldots \\ sym. & & & b_{kk} \end{bmatrix} \quad (2.21)$$

If $\hat{y}$ is differentiated respect to x,

$$\partial\hat{y}/\partial x = b + 2\hat{B}x \quad (2.22)$$

and this derivative is set to 0, the stationary point will be found as;

$$X_s = -B^{-1}b/2 \tag{2.23}$$

X_s is the stationary point of the system. The predicted response at the stationary point can be found by substituting equation (2.23) in to equation (2.18), it is;

$$\hat{y}_s = b_0 + x_s'b/2 \tag{2.24}$$

(Myers & Montgomery, 1995, p.218).

After finding stationary point, we determine whether the stationary point is a point of maximum response (Figure 2.5) or minimum response (Figure 2.6) or a saddle point (Figure 2.7). The most straightforward way to do this is to examine a contour plot of the fitted model. If there are only two or three process variables, the construction and interpretation of this contour plot is relatively easy.

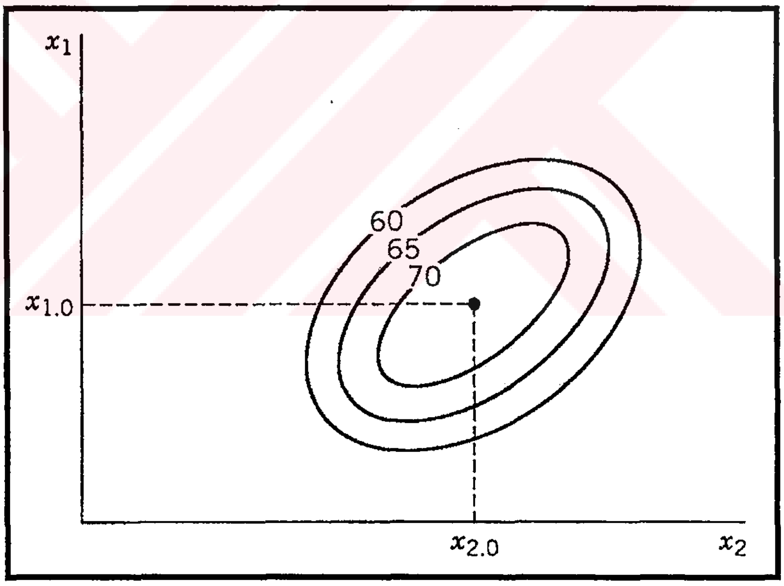

Figure 2.5 Stationary point is a point of maximum response

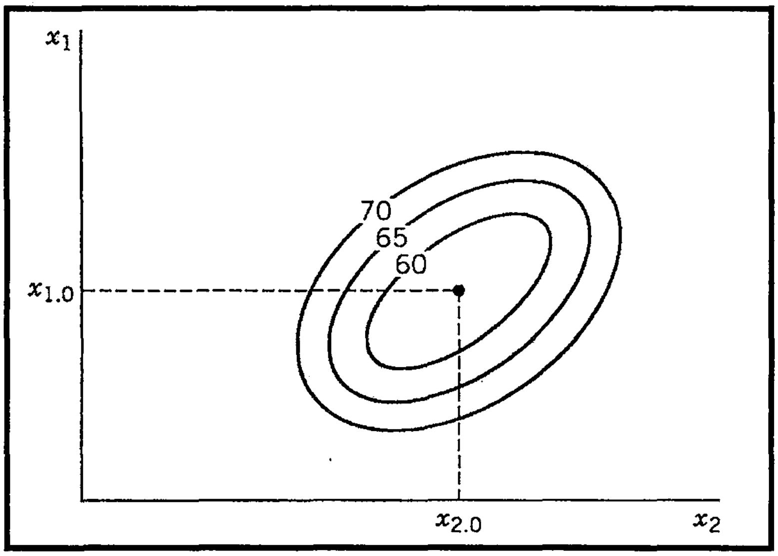

Figure 2.6 Stationary point is a point of minimum response

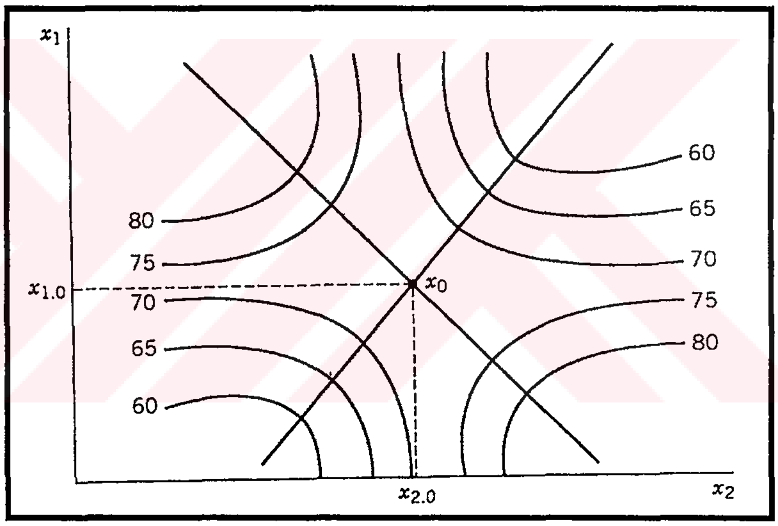

Figure 2.7 Stationary point is a saddle point

However, even when there are relatively few variables, a more formal analysis, called the canonical analysis, can be useful. It is helpful first to transform the model into a new coordinate system with the origin at the stationary point X_s and then to rotate the axes of this system until they are parallel to the principal axes of the fitted response surface. Results in the fitted model that is called canonical form of the model is;

$$\hat{y} = \hat{y}_s + \sum_{i=1}^{k} \lambda_i w_i^2 \tag{2.25}$$

where $\hat{y}_s$ is the estimated response at the stationary point, λ_i are the eigenvalues of $\hat{B}$ and w_i are canonical variables (transformed independent variables). Canonical form of the second-order model with two independent variables is denoted in Figure 2.8 (Montgomery, 2001, p.440).

Equation (2.25) nicely describes the nature of the stationary point and the nature of the system around stationary point. The nature of X_s (stationary point) is determined by the sings of the λ's;

- If λ_i are all negative, the stationary point is a point of maximum response.
- If λ_i are all positive, the stationary point is a point of minimum response.
- If λ_i are mixed in sign, the stationary point is a saddle point (Myers & Montgomery, 1995, p.219).

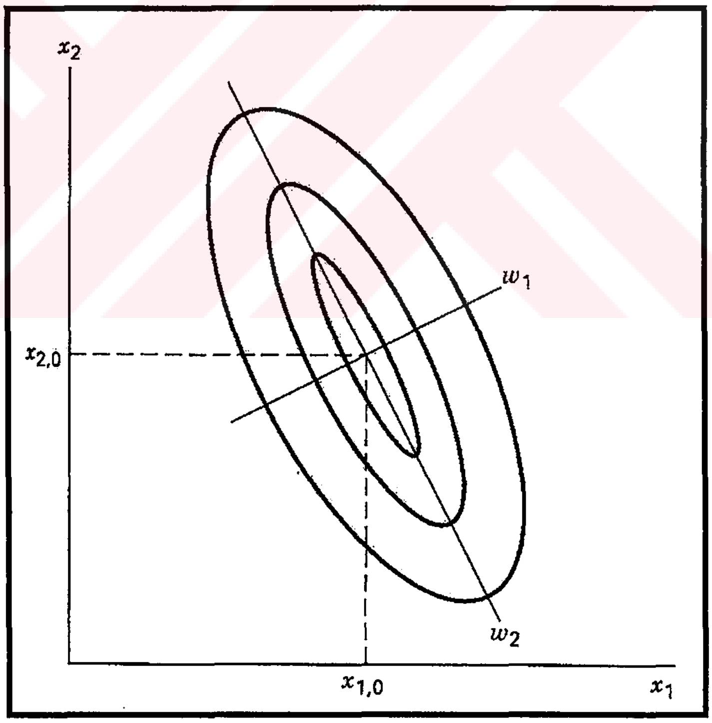

Figure 2.8 Canonical form of the second order model

It is not unusual to encounter variations of the pure maximum, minimum, or saddle point response surfaces, that system is called ridge system. The two type of ridge system are the stationary ridge system and the rising ridge system.

If the stationary point is in the region of the experimental design and one (or more) λ_i is near zero, the system is a stationary ridge system. As an example in Figure 2.9, a contour plot of a stationary ridge system with two independent variables is shown, and maximization is the objective of response. For this example λ_1<0 and λ_2<0 but $\lambda_1 \cong 0$, the canonical model for this response surface is theoretically;

$$\hat{y} = \hat{y}_s + \lambda_2 w_2^{\ 2} \tag{2.26}$$

clearly the stationary point is a point of maximum response, but there is essentially a line maximum. The response variable is very insensitive to the variable w_1 multiplied by the small λ_1. Optimum may be taken anywhere along w_1 axis.

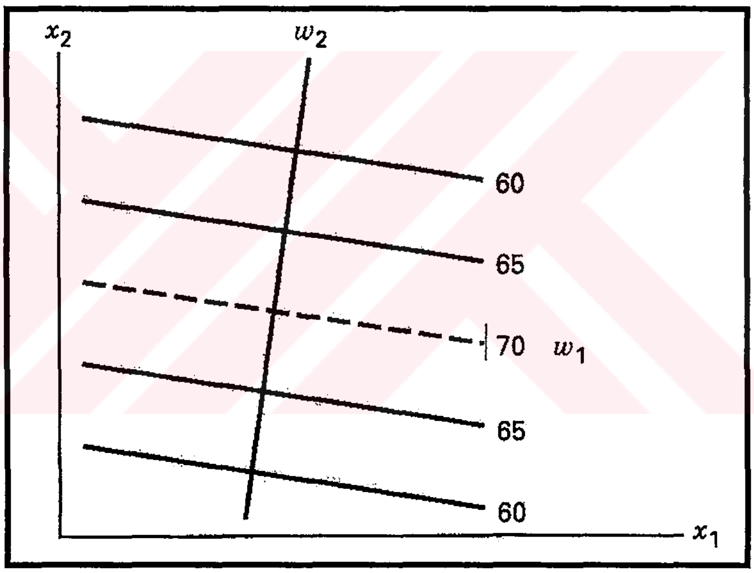

Figure 2.9 A contour plot of a stationary ridge system

If the stationary point is far outside the region of exploration for fitting the second order model and one (or more) λ_i is near zero, then the surface may be a rising ridge. In this type of ridge system, inferences about the true surface or the stationary point cannot be drawn because X_s is outside the region where the model fitted.

As an example in Figure 2.10, a contour plot of a rising ridge system with two independent variables is demonstrated, and again maximization is the objective of response.

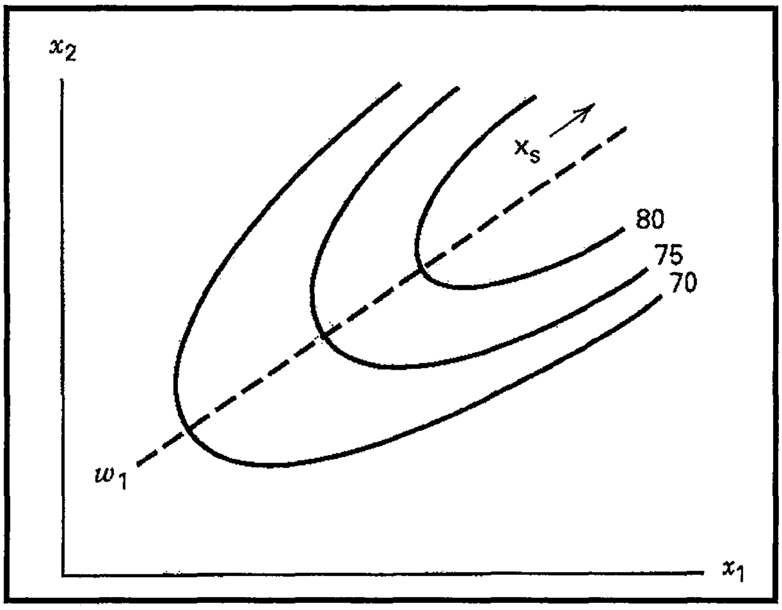

Figure 2.10 A contour plot of a rising ridge system

Again, $\lambda_1<0$ and $\lambda_2<0$ but $\lambda_1 \cong 0$, the canonical model for this response surface is same as in equation (2.26). Also it is seen that, further exploration is warranted in the w_1 direction. Indeed the rising ridge often is a signal to the researcher that he/she has perhaps made a faulty or premature selection of the experimental design region (Myers & Montgomery, 1995, pp.210-221) and (Montgomery, 2001, pp. 436-441).

2.7 Response Surface Analysis with Multiple Responses

Up to now, it is focused on modelling a measured response or a function of design variables and letting the analysis indicate areas in the design region where the process is likely to give desirable results, the term "desirable" being a function of the predicted response. However, in many instances the term desirable is a function of more than one response (Myers & Montgomery, 1995, pp.244).

Multiple response optimization is a method that allows for compromise among the various responses. Simultaneous consideration of multiple responses involves first building an appropriate response surface model for each response and then trying to find a set of operating conditions that in some sense optimizes all responses or at least keeps them in desired ranges.

A relatively straightforward approach to optimizing several responses that works well when there are only a few process variables is to overlay the contour plots for each response. The experimenter can visually examine the contour plot to determine appropriate operating conditions (Montgomery, 2001, pp. 448-450).

2.7.1 The Desirability Function

The overlaying of contour plots along with separate response surface analyses often give the user workable solutions for product improvement as long as the number of responses is not too great. However, when the problem involves four or more responses or design variables, then contour overlay methodology becomes unruly. Derringer & Suich (1980) developed an interesting procedure, which can be very useful when several responses are involved. The method makes use of a desirability function in which the researchers own priorities and desires on the response values are built into one optimization procedure (Myers & Montgomery, 1995, pp.247-248).

Engineers design products or processes by selecting $x_1, x_2, \ldots, x_p$ that will result in a desirable combination of properties or quality criteria, $Y_1, Y_2, \ldots, Y_m$. Functions that transform a set of properties into a single objective are called "desirability" functions and are written as $D(Y_1, Y_2, \ldots, Y_m)$ (Ribardo & Allen, 2003, p.227).

The desirability function approach is one of the most widely used methods in industry for the optimization of multiple response processes. It is based on the idea that the "quality" of a product or process that has multiple quality characteristics, with one of them outside of some "desired" limits, is completely unacceptable. The method finds operating conditions that provide the "most desirable" response values. The desirability approach is a popular method that assigns a "score" to set of responses and chooses factor settings that maximize that score (WEB_1, 2004).

The basic idea of the desirability function approach is to transform a multiple response problem in to a single response problem by means of mathematical transformations (Castillo et al., 1996, p.337).

First, Harrington (1965) introduced the concept of a "desirability function" for determining input parameter settings, which optimize the tradeoffs among multiple process measurements. Then, the procedure refined by Derringer & Suich (1980). In addition, a recent article by Castillo et al. (1996) suggested a slight modification to the procedure so that it can be easily implemented using the "Solver" function of Microsoft Excel (Fuller & Scherer, 1998, p.4016).

In Harrington (1965) exponential functional forms were selected to calculate the desirabilities associated with individual criteria, Y_i, and the use of the geometric mean for weighting these criteria together to calculate overall desirability. Derringer & Suich (1980) criticized the functional forms and weighting scheme in Harrington for being overly rigid. As an alternative, they suggested a family of functions that permitted the target value to be anywhere in the region between product specifications. Castillo et al. (1996) improved the individual criteria desirabilities of Derringer (1994) in order to achieve greater smoothness and differentiability. This smoothness is useful because it can improve the performance of gradient-based solvers in optimizing the derived desirability functions (Ribardo & Allen, 2003, p.228).

2.7.1.1 The Desirability Function of Harrington

Harrington used two steps to calculate the desirability function. The first step concentrated on each criterion/response to assign an individual desirability. Criteria divided into two types: 'two-sided' criteria whose acceptable values were bounded by both an upper specification limit (*USL*) and a lower specification limit (*LSL*) and 'one-sided' criteria whose acceptable values were bound by a single specification limit. For two-sided criteria, the process of assigning desirability by calculating the scaled response value $Y_j'(x)$ was performed using;

$$Y_j'(x) = \frac{2Y_j(x) - (USL + LSL)}{USL - LSL} \tag{2.27}$$

Then, the user would need to input a single scaled criterion value for criterion j, $Y_{j,0}$, and its assumed desirability d_0 (somehow independently of the values of other criteria), e.g. when $Y'_{j,0} = -0.1$ then d_0 = 0.63. This would be appropriate if −0.1 (slightly lower than the midpoint between the specification limits) corresponded to a 'good' system. This pair of numbers ($Y'_{j,0}$, d_0) was used to calculate the parameter n using the following formula;

$$n = \frac{\ln[\ln(1/d_0)]}{\ln|Y'_{j,0}|} \tag{2.28}$$

Next, the desirability for the two-sided characteristic was obtained using;

$$d_j(Y_j(x)) = \exp[-|Y_j'(x)|^n] \tag{2.29}$$

For single-sided criteria, the individual desirability measures or 'desirabilities' were calculated as follows. The engineer had to input two pairs $(Y_{j,1}, d_1)$ and $(Y_{j,2}, d_2)$ with the criteria values, $Y_{j,1}$ and $Y_{j,2}$, given in actual response units and it was assumed, without loss of generality, that $Y_{j,1} > Y_{j,2}$. Each of the response values was then scaled using the formula;

$$Y'_{j,1} = -\ln[-\ln(d_1)] \quad \text{and} \quad Y'_{j,2} = -\ln[-\ln(d_2)] \tag{2.30}$$

Then, the scaled criteria value, $Y_j'(x)$, corresponding to the actual response $Y_j(x)$ was found through the following linear transformation;

$$Y_j'(x) = [(Y_j(x) - Y_{j,2})/(Y_{j,1} - Y_{j,2})](Y'_{j,1}(x) - Y'_{j,2}(x)) + Y'_{j,2}(x) \tag{2.31}$$

Next, the desirability for the one-sided characteristic was estimated using;

$$d_j(Y_j(x)) = \exp[-\exp(-Y_j'(x))] \tag{2.32}$$

In the second stage to estimate the system desirability, individual criteria desirabilities were combined using the following formula involving subjective weights of each of the criteria, w_i;

$$D(x) = \left[d_1(Y_1(x))^{w_1} d_2(Y_2(x))^{w_2} \ldots d_m(Y_m(x))^{w_m}\right]^{1/S} \quad (2.33)$$

where $S = \Sigma_i \; w_i$.

In its original formulation, it was argued that $w_i = 1$ for all responses i was sufficient for most cases of interest. In subsequent research, several advantages for considering unequal weights were described in Derringer (1994). These included that weighting provides a more direct method to adjust the relative importance of alternative criteria than changing the other desirability parameters, e.g. Y_{j0} and d_0. Harrington's rating system for interpreting the desirability is denoted in Table 2.1.

Table 2.1 Harrington's rating system for interpreting the desirability, *d*

Rating	Description
1.00	The ultimate in satisfaction and quality (an improvement beyond this point would have no appreciable value)
1.00–0.80	Acceptable and excellent (represents unusual quality or performance well beyond anything commercially available)
0.80–0.63	Acceptable and good (represents an improvement over the best commercial quality)
0.63–0.40	Acceptable but poor (quality is acceptable to the specification limits but improvement is desired)
0.40–0.30	Borderline (if specification exists, then some of the product quality lies exactly on the specification maximum or minimum)
0.30–0.00	Unacceptable (materials of this quality would lead to failure)
0.00	Completely unacceptable

2.7.1.2 The Desirability Function of Derringer-Suich

One and two-sided desirability functions are used depending on whether the response is to be maximized or minimized or has an assigned target value. For a response with target value level, A, B, and C are assigned so that $A \leq B \leq C$. A product is considered unacceptable if $\hat{y} < A$ or $\hat{y} > C$. The value B is the "most desirable value" (target). The quantity d_i, the desirability, is defined as;

$$d_i = \begin{cases} \left(\dfrac{\hat{y}-A}{B-A}\right)^s, A \le \hat{y} \le B \\ \left(\dfrac{\hat{y}-C}{B-C}\right)^t, B \le \hat{y} \le C \end{cases} \tag{2.34}$$

with d being 0 if $\hat{y} > C$ or $\hat{y} < A$. One can then use Figure 2.11 to allocate power values s and t according to one's subjective impression about the role of this response in the total desirability of the product.

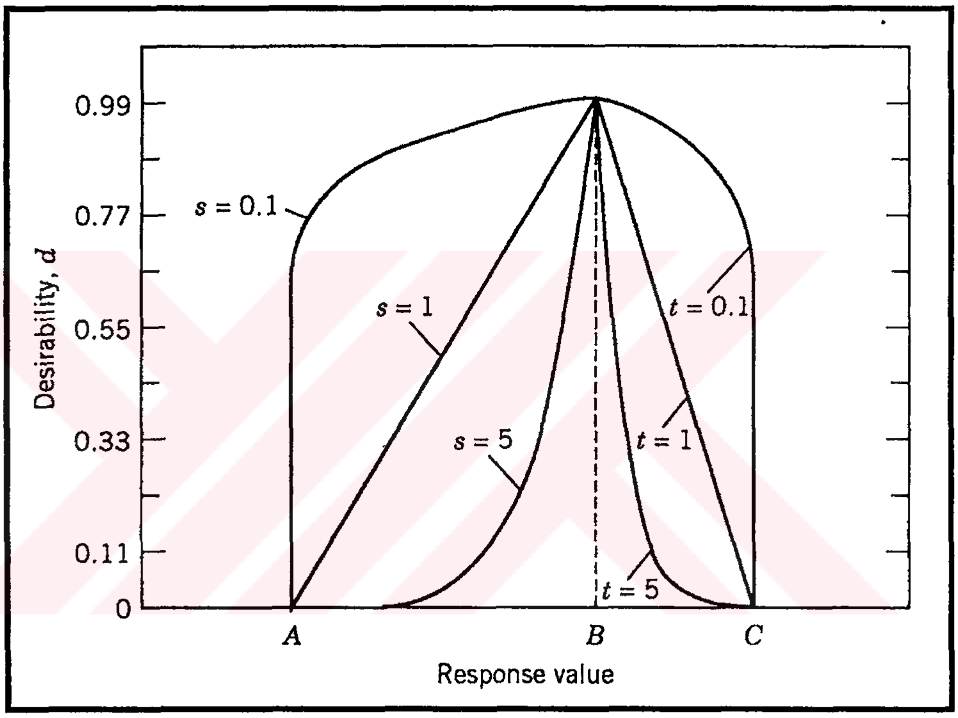

Figure 2.11 Desirability function for target value B

In the case where a response should be maximized, one chooses a value B such that $d = 1$ for any $\hat{y} > B$. However, we assume that the product is unacceptable (d = 0) if $\hat{y} < A$. Here, B = C, then the desirability function is given by;

$$d_i = \left\{ \left(\dfrac{\hat{y}-A}{B-A}\right)^s, A \le \hat{y} \le B \right. \tag{2.35}$$

In the case where a response is to be minimized, a value B is chosen such that $\hat{y} \le B$ is quite desirable and produces a d = 1. A value for $\hat{y} > C$ is considered unacceptable and therefore results in a $d = 0$. In this case, $A = B$. The desirability function is given by;

$$d_i = \left\{ \left(\frac{\hat{y} - C}{B - C} \right)^t, B \le \hat{y} \le C \right. \quad (2.36)$$

In most applications, values *A, B,* and *C* are chosen according to the researcher's priorities. Choices for *s* and *t* may be more difficult. Choice of *s* and *t* are determined by how important it is for $\hat{y}$ to be close to the target *B*. Using small values for *s* and *t* essentially does not require the response to be close to target, but a choice of *s* and *t* as large as implies that the desirability value is very low unless $\hat{y}$ gets very close to target For this reason, it can be tried for various levels of *s* and *t* (Myers & Montgomery, 1995, pp.250-251).

The general approach is to first convert each response y_i into an individual desirability function d_i that varies over the range $0 \le d_i \le 1$, where if the response y_i is at its goal or target, then $d_i = 1$, and if the response is outside an acceptable region, $d_i = 0$. Then the design variables are chosen to maximize the overall desirability;

$$D = (d_1 * d_2 * ... * d_m)^{1/m} \quad (2.37)$$

where there are *m* responses (Montgomery, 2001, pp. 451).

Although a given increase in $\hat{y}$ does not necessarily result in a uniform increase in the decision maker's preference, in the transformed desirability space, however, each increment in d_i represents the same marginal return to the decision maker (Fuller & Scherer, 1998, p.4017).

The overall desirability, *D*, is the geometric mean of the individual desirability values. This provides an overall assessment of the combined response functions levels. Maximizing *D* enables the simultaneous optimization of the geometric mean of the transformed response functions. The value of *D* that close to 1 implies that, all responses are in a desirable range simultaneously (Osborne et al., 1997, p.3834).

The rationale behind using the geometric mean in formula (2.33) and (2.37) is that if any quality characteristics has an undesirable value (i.e., $d_i = 0$) at some operating conditions, then the overall result is usually a product which is wholly unacceptable, regardless of the values taken on by the other responses (Castillo et al., 1996. p.337).

After the "optimum" or desirable condition on x has been determined, the researcher should do confirmatory runs at that condition to be sure that all responses are in a satisfactory region (Myers & Montgomery, 1995, p.252).

The appearance of the desirability function for the maximum $\hat{y}$ case, the impact of various s values are shown in Figure 2.12.

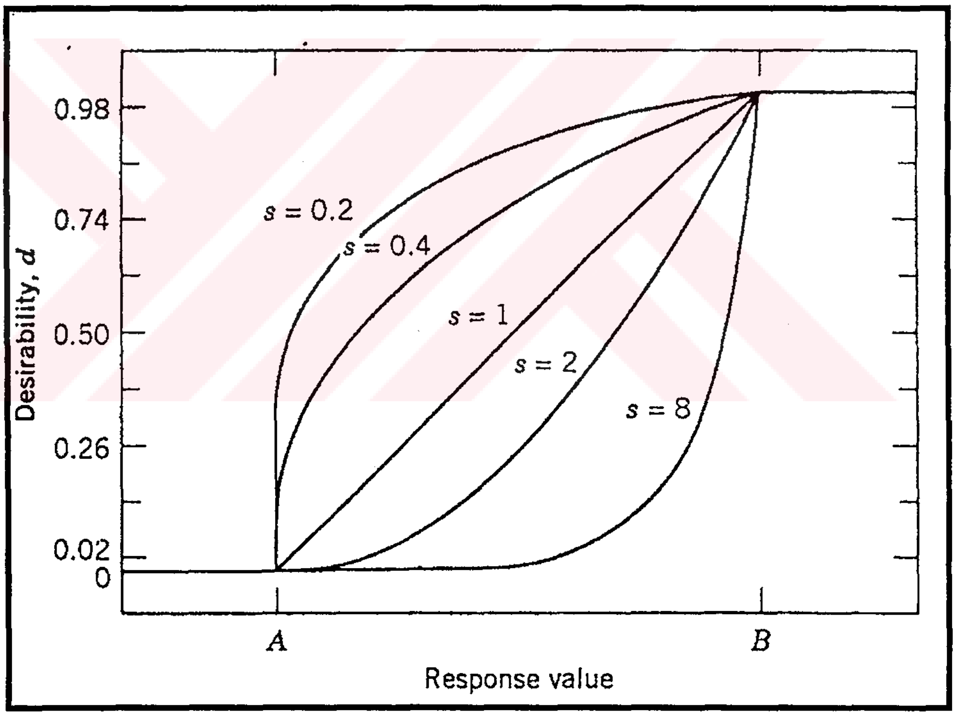

Figure 2.12 Desirability function for a response to be maximized

The decision maker may consider gains in the i^{th} response to be of constant, increasing or decreasing marginal worth. The rate of the marginal desirability of increases in the i^{th} response is controlled by the variable "s" in the transformation equation (2.35). If the marginal desirability is constant then there is a linear relationship between $\hat{y}$ and d_i and as such the value of s is set to 1. If additional units of $\hat{y}$ above the minimum acceptable value A are of decreasing marginal worth, then there is a concave relationship between $\hat{y}$ and d_i and as such the value of s is less than 1. Finally, if additional units of $\hat{y}$ above the minimum acceptable value A are of increasing marginal worth, then there is a convex relationship between $\hat{y}$ and d_i and as such the value of s is greater than 1. In the case that any value above the minimum is acceptable, a small value s would be selected (e.g. s = 0.2). In the case where approaching the process maximum, B, is considered highly desirable, then a large value of s would be selected (Fuller & Scherer, 1998, p.4017).

2.8 Literature Review on *RSM*

The literature can be divided into two main parts, first building a response surface metamodel and second optimizing the metamodel.

Building a Metamodel

Yu & Popplewell (1994) and Barton (1994) reviewed the literature and classified papers from different point of view. Yu & Popplewell (1994), reviewed the published development of metamodelling techniques, with particular emphasis on their relevance to manufacturing. They surveyed and categorized the papers which were published in from 1975 to 1993. They classified the papers into five class; *case studies* (that reported the application of metamodels), *general* (that exercised metamodelling techniques in hypothetical cases), *survey* (that surveyed the metamodelling tools or statistical tools associated with metamodelling), *technical* (that considered construction, implementation and validation of metamodels), and *others* (that were discussion papers or responses to previous reports). Also, Barton (1994) reviewed the literature and classified the metamodels from a different point of view such as, response surface metamodels, spline metamodels, radial basis function metamodels, kernel smoothing metamodels, and spatial correlation metamodels, and he mentioned mathematical form and experimental designs for these metamodels.

Some papers [Madu (1990), Sargent (1991), Groenendaal & Kleijnen (1996)] dealt with the technical points of metamodels such as, building phases of metamodel, and benefits obtained from metamodels. Madu (1990) illustrated the application of regression metamodels to simulation outputs on a maintenance float problem. He used the metamodel to conduct sensitivity analysis, also optimization and validation of the metamodel were demonstrated. He proposed a series of steps to be used in metamodel development and also explained common benefits of the metamodel. Also, Sargent (1991) after giving an overview of the metamodelling process, identified and discussed some of the research issues in metamodelling of simulation models. As an example, metamodelling of the M/M/1 queues is used for illustrating some of these issues. In addition, Groenendaal & Kleijnen (1996) thought that design of experiments in combination with regression metamodelling can be a fruitful method for problems with uncertainties that do not fit into the standard uncertainty models and that have no reliable information on the joint probability distribution. They also illustrated design of experiments and regression analysis by a large practical investment problem.

Some authors focused on validation of metamodels as Kleijnen & Sargent (2000), who derived a methodology for developing metamodels that considers validation of a metamodel with respect to both the underlying simulation model and the problem entity. In methodology they distinguished between fitting and validating, emphasized the role of the problem entity, and also distinguished four types of goals such as understanding, prediction, optimization, and verification & validation. They also gave a detailed procedure for validating linear-regression metamodels in random simulation, suitable to all four goals. Also, Santos & Nova (1999) proposed a methodology for identifying a tentative functional relationship, estimating the nonlinear simulation metamodel coefficients and validating the metamodel.

Some studies aimed to build decision support systems by using metamodels. McHaney & Douglas (1997) demonstrated how a representational decision support system (*DSS*) can be simplified into a suggestion model *DSS* without a loss of accuracy. This transformation was accomplished by using a multivariate regression metamodel developed from a discrete event simulation model to a linear equation. They used simulation metamodel in an automotive assembly plant environment to assist management in determining automated guided vehicle requirements. Production management has the ability to answer resource allocation question rapidly and efficiently by placing the resulting equation into a small computer program.

In addition, metamodels were used for supporting the evaluation of the manufacturing systems. Sridharan & Babu (1998) pertained to a detailed simulation study conducted on a typical flexible manufacturing system (*FMS*). Simulation models had been developed for two types of *FMS*s, a failure free and a failure prone and the simulation results used to develop metamodels. These metamodels were used to evaluate various multi-level scheduling decisions in the *FMS*. And, Jothiskankar & Wang (1993) demonstrated application of simulation metamodelling for a just-in-time (*JIT*) manufacturing environment. They illustrated that satisfactory prediction for throughput time can be obtained through regression metamodels. In addition, Aytuğ et al. (1996) analyzed a pull production system by using simulation metamodel in order to determine the number of kanbans. A relationship between the number of kanbans and the average time to fill a customer order was established by the metamodelling process. They used the relationship in a cost function to determine the number of kanbans while minimizing cost.

Optimizing the Metamodel

The literature about *RSM* can be classified as surveys, technical papers that dealt with the theory of methodology and designed experiments, case studies that explained the applications of methodology in manufacturing environmental, and also studies that dealt with multiple responses.

The newest survey study is Myers et al. (2004) which reviewed the *RSM* studies that had been published since 1989, and also discussed current areas of research and mentioned some areas for future research.

Two papers [Brightman (1978), Hood & Welch (1993)] were aimed to present the methodology in detail. Brightman (1978) presented the all phases of the *RSM* with a numerical example in order to explain all aspects of the methodology. And, Hood & Welch (1993) presented an outline of the response surface methodology and gave an example of its application to a two variable optimization problem. Another technical paper is Angün et al. (2002) which derived adopted steepest descent (*ASD*) which is scale independent and corrects for the covariance of the estimated gradient components because of the scale dependence nature of steepest descent (*SD*). Also, Batmaz & Tunali (2003) aimed to provide guidance on how to use small designs for metamodel estimation especially when cost effectiveness is a concern. Their study was carried out in three phases: First, a group of second order small designs were evaluated with respect to various criteria. Second, the metamodel of a time-shared computer system was estimated using these designs. Third, the predictive capabilities of these small designs in giving the best metamodel fits were investigated, and also, the performance of small designs were compared with two large size standard designs.

The following papers dealt with the applications of methodology in manufacturing environmental. Mahadeevan & Narendran (1993) addressed the issues involved in the choice of material handling systems and the buffer capacity of *FMS*s with medium and high congestion. In order to determine the capacities of local and central buffer that minimize the throughput time in two environments, they employed simulation and *RSM*. Another paper relevant to *FMS* D'Angelo et al. (1998), which focused on the typical job-shop plant configuration for the semiautomatic manufacturing of parts produced in limited quantities. The treatment consists of quantitative evaluation of the technological performance of a *FMS* with particular reference to the printed circuit board assembly sector. *RSM* is applied for balancing the capacities of the work centres of a printed circuit board assembly leading industry. In addition, Gharbi & Kenne (2000) presented a production control approach for manufacturing systems that is based on simulation experiment. The objective of the control problem was finding the production and preventive maintenance rates of the machines so as to minimize the total cost of inventory. By combining analytical formalism and simulation based statistical tools such as experimental design and *RSM* an approximation of the optimal control policy is obtained. And, Irizarry et al. (2001a) presented a general manufacturing-cell simulation model for evaluating the effects of world-class manufacturing practices on expected cell performance. The modular structure of the simulation model provided the flexibility to analyze a wide variety of manufacturing cells. They formulated a comprehensive annualized cost function for evaluation and comparison of alternative cell configurations. In addition, a case study involving assembly of printed circuit boards illustrated the potential benefits of using this tool for cell design and analysis. The simulation model was intended for use in a two-phase approach to cell design that is based on simulated experimentation and response surface analysis as detailed in Irizarry et al. (2001b). In the companion paper, Irizarry et al. (2001b) constructed simulation (response surface) metamodels to describe the relationship between the significant cell design operational factors (the controllable input parameters) and the resulting simulation-based estimate of expected annual cell cost (the output response). They used canonical and ridge analyses of the estimated response surface to estimate the levels of the quantitative input factors that minimize the cell's expected annual cost. And, Horng & Cochran (2001) presented a decision support methodology called project surface regions (*PSR*). The methodology developed via discrete event simulation and *RSM* and utilized the usage of *PSRs* to assist production manager in determining the appropriate number of multitasking workers and the corresponding dynamic dispatching rule-pair when *JIT* system's behaviour changes.

In some circumstances the response can be affected by both qualitative and quantitative factors as in Wu & Ding (1998), Irizarry et al. (2001b), and Tunali & Batmaz (2003). Wu & Ding (1998) proposed a general approach for constructing response surface designs of economical size for qualitative and quantitative factors. Algorithm starts with an efficient design for the quantitative factors and then partitions the design points into groups corresponding to different level combinations of the qualitative factors. Also, Tunali & Batmaz (2003) suggested a methodology for developing a simulation metamodel involving both quantitative and qualitative factors. The methodology dealt with various strategic issues involved in metamodel estimation, analysis, comparison, and validation. To illustrate how to apply the methodology, a regression metamodel is developed for a client–server computer system.

Some papers dealt with multiple responses and also desirability functions. Mollaghasemi et al. (1991) reviewed some available techniques for solving multi-response simulation model, and also proposed a new method that was based on the gradient search technique. Osborne et al. (1997), described the available methods for optimizing multiple response variables using *RSM*, and several schemes for categorizing the methodologies were delineated. The applications of multiple response surface methodologies (*MRSM*) in the area of product development, and also the use of simulation to assist in the applications of *MRSM* in the product development context were described. Another study was Leon & Cabrera (1997), which dealt with the effective use of experimental design techniques for assessing and optimizing process capability in the manufacturing of electrical igniters. A factorial design was used to assess the effects that four factors have on five responses. An approach harmonizing traditional specification limits with current notions of loss functions was used for settings that simultaneously optimize the response variables. And, Myers et al. (2004) classified the approaches about multiple response optimization into four classes such as, contour plots approach, formulating the multi response problem as a constrained optimization problem approach, simultaneously optimizing all *m* responses approach, and the desirability function approach.

The desirability function concept was first introduced by Harrington (1965) for determining input parameter settings, which optimize the tradeoffs among multiple process measurements. Then, the procedure refined by Derringer & Suich (1980). Also, Derringer (1994) dealt with balance problems that arise because as one property is improved, it is often at the expense of one or more other properties, and also declared balance problems can be solved by using a modified formula of Harrington's desirability function and combining it with *RSM* to form a methodology called desirability optimization methodology. Computer implementation of desirability optimization methodology further enhances its power. Because the original formulation of desirability functions contain non-differentiable points, only search methods can be used to optimize the overall desirability response, thus Castillo et al. (1996) presented modified desirability functions which were everywhere differentiable, so that more efficient gradient-based optimization methods can be used. In addition, they suggested a slight modification to the procedure so that it can be easily implemented by using the "Solver" function of Microsoft Excel. And also, Fuller & Scherer (1998) thought that little attention had been paid to the underlying assumptions of decision theory on which the desirability function procedure is based. Therefore, they demonstrated critical issues that can arise from this procedure. Another study that used desirability functions were Dabbas et al. (2001) that validated a modified dispatching approach, which is proposed in Dabbas & Fowler (1999), and which combines multiple dispatching criteria into a single rule with the objective of simultaneously optimizing multiple objectives. Dabbas et al. (2001) used the Derringer & Suich's desirability function for optimizing multiple responses. In addition, Ribardo & Allen (2003) proposed a new desirability function to aid in multi-criterion optimization. They declared the advantages of the proposed function as; first, its value has the simple interpretation of being the yield under conservative assumptions that are standard in the six sigma literature. Second, because the proposed function is expressed explicitly as a function of the mean and standard deviation of the associated criteria or quality characteristics, the settings derived using this function are more likely to correspond to the true preferences of engineers than if existing alternative desirability functions are used.

2.9 Conclusion

RSM is the body of techniques by which one seeks on optimum set of operating conditions. For a single response model, the methodology generally consists of constructing an appropriate first order experimental design such as 2^n factorial or 2^{n-p} fractional factorial designs around an initial feasible point x^k. A simulation trial is then performed at each design point and the observations are recorded. At the initial stages, since the point selected is very likely to be far from the optimal solution, a first order model is fit through the points by method of least squares. The method of steepest ascent can then be employed to search for a better solution along the direction of steepest increase of the response. The search continues until the first order equation can no longer improve the response function. This condition can be detected by a lack-of-fit test. When this situation is encountered, the search process is probably close to an optimal point. At this point a second order experimental design such as central composite design is constructed and a second order response equation is generated by regression analysis. Once this second order regression equation is estimated, canonical analysis can be employed to obtain an optimal solution (Mollaghasemi et al., 1991,p.201).

When multiple responses are considered simultaneously, it is first built an appropriate response surface model for each response and then, tried to find a set of operating conditions that in some sense optimizes all responses or at least keeps them in desired ranges (Montgomery, 2001, p. 448).

CHAPTER THREE
PROBLEM DEFINITION

3.1 Introduction

Because of the increasing car traffic and the need to improve the environment of the Izmir, Turkey, a public metro-line has been constructed. This underground metro-line, began working full time in August 26, 2000. On the line, there are ten stations that trains running. Four of them; Ucyol, Konak, Cankaya and Basmane are under ground, two of them; Hilal and Stadyum are on viaduct, three of them; Halkapinar, Sanayi and Bolge are on the ground, and Bornova is in unroofed slot tunnel. Ucyol, Hilal and Stadyum stations are side type, and the others are land type stations. The map of Izmir City centre and the present metro line is shown in Figure 3.1.

There are ten working time periods in a week from Monday to Friday, four working time periods in Saturday and three working time periods in Sunday. The starting and ending time of the periods, and related headways are shown in Table 3.1.

Each time period has its own headway that is the time period between the departure times of two consecutive trains, and is fixed in a certain time period. Headway determines the number of trains in a certain period, and the number of trips in a day depends on headways.

The metro company has to make attractive the use of transport system for passengers. Therefore, the problem is to find the headways (input factors), for each ten time interval in five days between Monday and Friday, for each four time interval in Saturday, and for each three time interval for Sunday. The objective is to minimize the average passenger time spent in the metro-line (the first response) with the requirement that the fullness rate of the carriages as fifty percent (the second response). We solve this problem by integrating *RSM* into simulation.

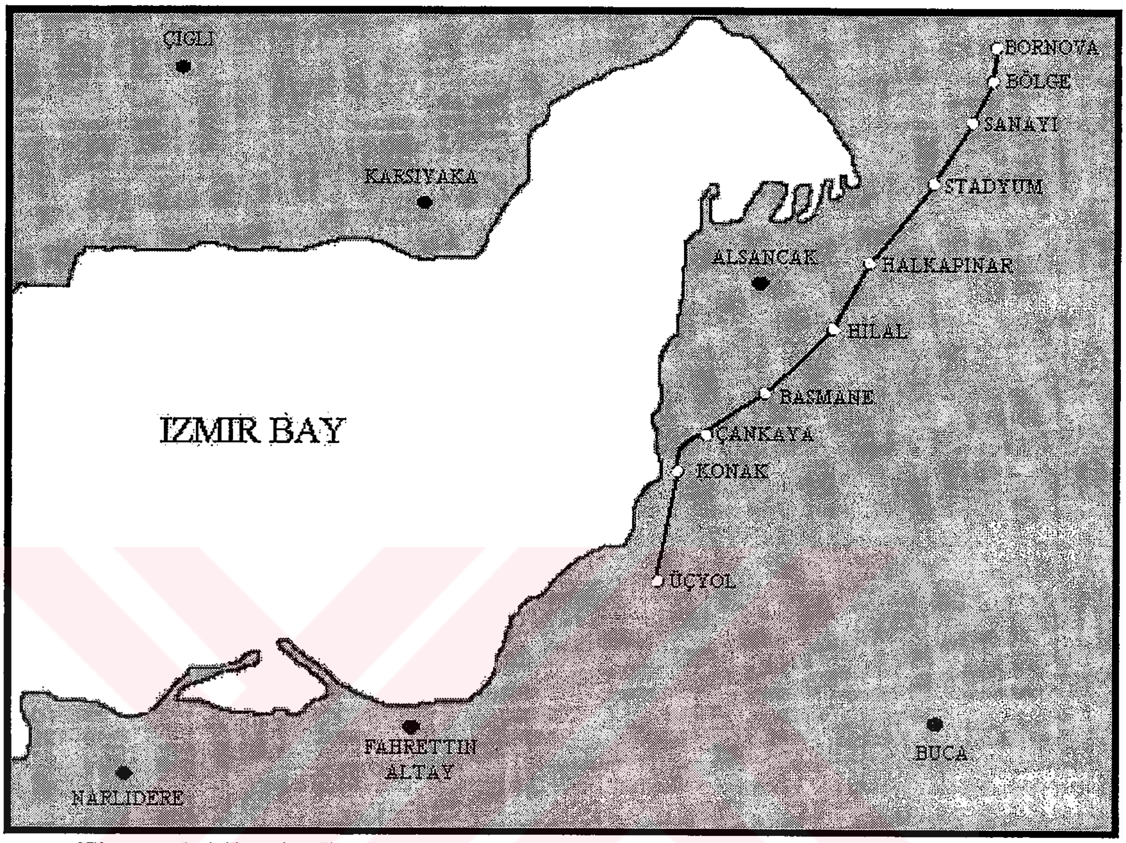

Figure 3.1 Izmir City centre map and present metro line with stations

Table 3.1 Time periods and current headways

Monday-Friday			Saturday		
Time period	Headway (minute)	Headway (second)	Time period	Headway (minute)	Headway (second)
06:00 - 07:00	10	600	06:00 - 11:00	10	600
07:00 - 07:30	7,5	450	11:00 - 19:00	7,5	450
07:30 - 09:00	5	300	19:00 - 22:00	10	600
09:00 - 09:30	7,5	450	22:00 - 00:00	15	900
09:30 - 11:30	10	600	**Sunday**		
11:30 - 17:00	7,5	450	Time period	Headway (minute)	Headway (second)
17:00 - 18:30	5	300			
18:30 - 19:00	7,5	450	06:00 - 09:00	15	900
19:00 - 22:00	10	600	09:00 - 20:00	10	600
22:00 - 00:00	15	900	20:00 - 00:00	15	900

It is obvious that the time spent in the system by a passenger directly depends on headways. However, while headways in a certain time period are decreased the number of empty carriages increases. Due to this fact, fullness rate of the carriages must be taken into consideration. Fullness rate is calculated by the following formula, and it is expected to be 50% by the company engineers.

$$FR = \frac{TNTP}{2*(TNT*564)} \tag{3.1}$$

where *FR* is the fullness rate, *TNTP* is the total number of travelled person in a day, *TNT* is the total number of trips in a day, 564 is the full capacity of a train that has 3 carriages, and 2 indicates there are two end stations. We assumed that all trains have 3 carriages.

3.2 System Definition

Bornova and Ucyol stations are the end stations, where trains begin to trip, and the other eight stations are the middle stations, where passengers wait train. Also, each day, when the metro system begin to serve, some middle stations behave as end station according to the number of trains (beginning trains) needed by the first time period. The distances between train stations are given in Table 3.2.

Table 3.2 The distances (in meter) between stations

Station	Bornova	Bolge	Sanayi	Stadyum	Halkapinar
Bornova	0	919	1989	2930	4157
Bolge	919	0	1070	2011	3238
Sanayi	1989	1070	0	941	2168
Stadyum	2930	2011	941	0	1227
Halkapinar	4157	3238	2168	1227	0
Hilal	5652	4733	3663	2722	1495
Basmane	6552	5633	4563	3622	2395
Cankaya	7132	6213	5143	4202	2975
Konak	8102	7183	6113	5172	3945
Ucyol	9509	8590	7520	6579	5352

Table 3.2 ***(Continued)***

Station	Hilal	Basmane	Cankaya	Konak	Ucyol
Bornova	5652	6552	7132	8102	9509
Bolge	4733	5633	6213	7183	8590
Sanayi	3663	4563	5143	6113	7520
Stadyum	2722	3622	4202	5172	6579
Halkapinar	1495	2395	2975	3945	5352
Hilal	0	900	1480	2450	3857
Basmane	900	0	580	1550	2957
Cankaya	1480	580	0	970	2377
Konak	2450	1550	970	0	1407
Ucyol	3857	2957	2377	1407	0

Each station, except Halkapinar, has two platforms, one of them is for train going from Ucyol to Bornova, and the other one is for train going from Bornova to Ucyol. In this thesis we let platform denote a segment of track on which a train stops at a station and only one train at a time is permitted. Halkapinar station wherein trains arrive and depart the system has three platforms. In Table 3.3 the platform lengths for each station are demonstrated.

Table 3.3 Platform lengths (in meter)

Station	Platform length	Station	Platform length
Bornova	135	Hilal	135
Bolge	135	Basmane	135
Sanayi	135	Cankaya	139
Stadyum	135	Konak	137
Halkapinar	135	Ucyol	137

There are two tracks between the two end stations, one of them is for travelling from Bornova to Ucyol and the other one is for travelling from Ucyol to Bornova. These two tracks are parallel to each other, and at some points they are connected to each other by switches. Thus, trains can change the track while they are travelling between stations.

Also there is a park area near Halkapinar station for empty carriages. For each time period the number of trains in the system changes due to headway. Thus, the number of carriages changes due to the number of trains in the system. Unused carriages wait at the park area, and when a new train is needed by the system, they are connected to each other to form a train.

Since the traffic flow is managed by the computer in the metro company, the tracks and switches that trains will follow are determined automatically. Interlocking and Automatic Train Protection (*ATP*) systems guarantee the travel safety. Electronically interlocking systems check and direct line-along equipments such as switches, signals, line circuits that facilitate remote control of trains for safely working.

Signalization system in Izmir Metro based on static block logic. Static block system separates a track into independent parts. Only one train may be in a part of track. Lengths and the number of blocks were determined by engineers related to traffic capacity of the system and speed of the trains.

3.3 Stations

In an end station, the train waits until the departure time. If there is no train at departure time passengers go on waiting train, and when the train arrives, passenger board and trip begins without losing any time.

In the middle stations, passengers wait for train. When the train arrives passenger board and trip goes on. It only waits for a period (dwell time) less than one minute for passengers alighting or boarding. In Table 3.4 dwell times, waiting times of trains in the middle stations, are given.

Halkapinar station, which is nearly at the middle of the metro-line, is somewhat different from other middle stations. Although it is a middle station it behaves as an end station at the beginning of each day, and trains both arriving and departing the system use this station. Another difference is that this station has three platforms although the other stations have two platforms. Because of these dissimilarities and more complex structure, Halkapinar station is explained in detail in the latter sections.

Table 3.4 Dwell times of trains at stations (in second)

Station	Dwell time	Station	Dwell time
Bolge	20	Hilal	20
Sanayi	20	Basmane	30
Stadyum	20	Cankaya	30
Halkapinar	20	Konak	40

3.4 Carriages

Carriages are designed as light rail vehicle type, and are in an accordion coupling unit shape with six axles and three bogies. There are forty-five carriages in the vehicle fleet. Thirty of them are with driver cabinet (*MD*), and the other fifteen are without driver cabinet (*M*). A train consists of three (*MD-M-MD*) or five (*MD-M-M-M-MD*) combinations of carriages. A carriage includes 44 seats and 36 square meter empty spaces where passengers can stand. The allowed number of standing passenger per meter square is four. A carriage has eight doors which are 1.4 meters in width.

The train's speed is limited. On the control panel, driver can see the speed limit on the current line, speed limit on the next line, status of signals, and status of switches. The *ATP* system warns the driver when permitted speed limit is exceeded, and also stops train if a risky situation occurs.

CHAPTER FOUR
SIMULATION STUDY

4.1 Features of the System

The system is a terminating system; it is opened at 05:45:00 a.m. and closed at 00:40:00 a.m.. thus, the system works totally 18 hours and 55 minutes (68100 seconds) in a day. But the first trip is at 06:00:00 a.m. and the last trip is at 24:00:00 a.m. Namely, the system is prepared before the first time period and it takes fifteen minutes. In a day between Monday and Friday, there are ten-time periods, which have their own headways. The number of time periods is four on Saturday, and is three on Sunday. Company engineers determine the lengths of these periods. At the beginning of a day, the trains (beginning trains) waiting in the park area enter the system from Halkapinar station, 15 minutes before the first time period begins. They go suitable stations, locate on determined platforms and waits for the starting time of the first time period.

4.1.1 Calculation of the Number of Trains Needed

Though the number of trains in the system is static in each time period, it is dynamic through a day. The number of trains needed for a time period depends on the headway in that time period, and is calculated by using formula (4.1).

$$NONT = \frac{RTT}{Headway} \tag{4.1}$$

where *NONT* is the number of trains needed for a time period, *RTT* is Round Trip Time, and *Headway* is the headway of related time period. *RTT* means the duration needed for a whole tour in the system, and it was as 40 minutes.

Since the headways can take values between 5 minutes and 15 minutes, according to the formula (4.1) the number of trains can be 3, 4, 5, 6, 7, and 8. For instance, if in a time period headway is 5 minutes then 8 trains (40 / 5 = 8) are needed, or if headway is 15 minutes then 3 trains (40 / 15 = 2.67) are needed.

As mentioned before, trains take certain positions on the platforms before the first time period begins. After that these trains will be called *beginning trains*. These positions are determined by this logic; if the number of initial trains is 3, a train will be located in Ucyol-Bornova direction (on platform 2), and the others in Bornova-Ucyol direction (on platform 1), and in Halkapinar- Ucyol direction (on platform 1), respectively. If the number of trains is 4, the three trains' locations are like the three trains' explained before. The fourth one will be in Halkapinar-Bornova direction (on platform 2). If 5 trains are needed, in addition to the previous locations, the fifth train will be in Konak- Ucyol direction (on platform 1). Using the same logic, location, platform number and direction of the initial trains are displayed in Table 4.1.

Table 4.1 Location, platform no and direction of the beginning trains

The beginning trains no	Location	Platform no	Direction
1	Ucyol	2	Bornova
2	Bornova	1	Ucyol
3	Halkapinar	1	Ucyol
4	Halkapinar	2	Bornova
5	Konak	1	Ucyol
6	Bolge	2	Bornova
7	Basmane	1	Ucyol
8	Stadyum	2	Bornova

4.1.2 Passenger Arrivals and Departures

The passengers arrive the system 5 minutes before the first time period begins. At 06:00 o'clock the beginning trains start to run. Passengers may enter the system until 23:59:30, that is 30 seconds before the last time period finishes. After that time no passenger is allowed to enter the system. Last trips begin at 24:00 and trains go on travelling until they reach the end station on their travelling direction. After passengers' departure, they turn to Halkapinar station without stopping, depart the system, and then locate on the park area.

4.1.3 Changing the Number of Travelling Trains

In the system, the number of running trains is adjusted due to the changes in headways. Therefore, the number of trains required in the next time period is calculated 10 minutes before the current time period finishes.

If the next time period has longer headway, the number of trains must be decreased. Therefore, the unneeded trains leave the system from Halkapinar station. Leaving action begins by the beginning of next period. For instance, an extreme condition may cause an increase in headway to 15 minutes by 10. That means the number of trains will be decreased to 3 by 5. That is, 5 trains will leave the system. Thus, the first train coming to the Halkapinar station, whatever direction it has, unloads its passengers, and leaves the system by using switches and then parks at the park area. This action goes on until 5 trains leave the system. Unloaded passengers wait a new train to go on their trip. If Halkapinar station is an unloaded passenger's destination he/she leaves the system without any delay.

If the next time period has shorter headway, the number of trains will be increased. New trains enter the system from Halkapinar station. For instance, if an extreme condition causes a decrease in headway to 5 minutes by 10 minutes, the number of trains will be increased to 8 by 5 for the next time period. 10 minutes before the current time period finishes, 2 trains enter the system from Halkapinar station. One of them is in Ucyol direction, and the other one is in Bornova direction. At that time the number of trains is increased to 5 by 2. But the system still needs 3 trains. 5 minutes later, that is the next period's headway, again 2 trains enter the system, towards Ucyol and Bornova, respectively. Now the number of trains is 7, yet less than 8, and then 5 minutes later the last needed train enters the system towards the Ucyol station. Passengers waiting a train at Halkapinar station may board these new trains.

4.2 Inputs of the Simulation Model

4.2.1 Passenger Arrivals to the Stations

Input Analyser module of Arena V2.2 simulation software package is used to find the distributions of passenger arrivals to each station for each time period. We used 5% (α = 0.05) significance level for chi-squared goodness of fit tests for arrivals. Observations on arrivals were given by the metro company, and are related to the last week of February 2004.

The distributions of arrivals in weekdays are given in Tables 4.2 - 4.6. The distribution values are in second unit.

Table 4.2 Arrivals to Bornova and Bolge on weekdays

Time period	Bornova	Bolge
05:55:00 - 06:00:00	-0.001 + WEIB(23, 0.679)	0.999 + WEIB(50.1, 0.752)
06:00:00 - 07:00:00	-0.001 + WEIB(23, 0.679)	0.999 + WEIB(50.1, 0.752)
07:00:00 - 07:30:00	-0.5 + EXPO(8.86)	-0.001 + WEIB(16.8, 0.861)
07:30:00 - 09:00:00	-0.5 + WEIB(3.69, 1.57)	-0.5 + WEIB(8.02, 1.35)
09:00:00 - 09:30:00	-0.5 + WEIB(4.25, 1.4)	-0.5 + WEIB(10.3, 1.35)
09:30:00 - 11:30:00	-0.5 + WEIB(5.24, 1.4)	-0.5 + WEIB(13.1, 1.25)
11:30:00 - 17:00:00	-0.5 + WEIB(4.01, 1.39)	-0.5 + EXPO(15.2)
17:00:00 - 18:30:00	-0.5 + WEIB(2.42, 1.38)	-0.5 + EXPO(14.1)
18:30:00 - 19:00:00	-0.5 + EXPO(3.88)	-0.001 + EXPO(11.7)
19:00:00 - 22:00:00	-0.5 + WEIB(5.66, 1.3)	-0.001 + EXPO(19.8)
22:00:00 - 23:59:30	-0.5 + EXPO(13.8)	-0.001 + WEIB(50.1, 0.73)

Table 4.3 Arrivals to Sanayi and Stadyum on weekdays

Time period	Sanayi	Stadyum
05:55:00 - 06:00:00	2 + EXPO(376)	-0.001 + WEIB(71.6, 0.846)
06:00:00 - 07:00:00	2 + EXPO(376)	-0.001 + WEIB(71.6, 0.846)
07:00:00 - 07:30:00	2 + EXPO(82.7)	-0.001 + EXPO(24.8)
07:30:00 - 09:00:00	-0.001 + EXPO(27.1)	-0.5 + EXPO(9.22)
09:00:00 - 09:30:00	-0.001 + EXPO(34.2)	-0.5 + EXPO(11.5)
09:30:00 - 11:30:00	-0.001 + WEIB(38.2, 0.873)	-0.5 + EXPO(14.6)
11:30:00 - 17:00:00	0.999 + EXPO(34)	-0.5 + EXPO(6.73)
17:00:00 - 18:30:00	-0.5 + EXPO(17.9)	-0.5 + EXPO(7.32)
18:30:00 - 19:00:00	-0.001 + EXPO(14.7)	-0.5 + WEIB(7.7, 1.16)
19:00:00 - 22:00:00	-0.001 + WEIB(28.1, 0.791)	-0.5 + EXPO(10.6)
22:00:00 - 23:59:30	-0.001 + WEIB(118, 0.731)	-0.001 + WEIB(70, 0.763)

Table 4.4 Arrivals to Halkapinar and Hilal on weekdays

Time period	Halkapinar	Hilal
05:55:00 - 06:00:00	2 + WEIB(102, 0.652)	0.999 + EXPO(248)
06:00:00 - 07:00:00	2 + WEIB(102, 0.652)	0.999 + EXPO(248)
07:00:00 - 07:30:00	0.999 + WEIB(70.9, 0.775)	0.999 + EXPO(87.1)
07:30:00 - 09:00:00	-0.5 + EXPO(15.5)	-0.001 + WEIB(31.1, 0.893)
09:00:00 - 09:30:00	-0.001 + WEIB(27.9, 0.862)	-0.001 + EXPO(58.2)
09:30:00 - 11:30:00	-0.001 + WEIB(30.9, 0.791)	-0.001 + EXPO(74)
11:30:00 - 17:00:00	0.999 + WEIB(21.9, 0.751)	-0.001 + EXPO(47.3)
17:00:00 - 18:30:00	-0.001 + WEIB(10.1, 0.716)	-0.001 + WEIB(33.8, 0.676)
18:30:00 - 19:00:00	-0.5 + WEIB(6.62, 1.23)	2 + EXPO(62.4)
19:00:00 - 22:00:00	-0.001 + WEIB(17.6, 0.669)	-0.001 + WEIB(117, 0.675)
22:00:00 - 23:59:30	0.999 + WEIB(155, 0.692)	18 + EXPO(461)

Table 4.5 Arrivals to Basmane and Cankaya on weekdays

Time period	Basmane	Cankaya
05:55:00 - 06:00:00	-0.001 + WEIB(58.1, 0.765)	-0.001 + WEIB(69.3, 0.704)
06:00:00 - 07:00:00	-0.001 + WEIB(58.1, 0.765)	-0.001 + WEIB(69.3, 0.704)
07:00:00 - 07:30:00	-0.001 + WEIB(23.7, 0.868)	-0.001 + WEIB(29.6, 0.852)
07:30:00 - 09:00:00	-0.001 + WEIB(7.94, 0.729)	-0.001 + WEIB(14.7, 0.898)
09:00:00 - 09:30:00	-0.001 + EXPO(22.8)	-0.001 + EXPO(13.2)
09:30:00 - 11:30:00	-0.001 + WEIB(20.3, 0.878)	-0.5 + EXPO(13)
11:30:00 - 17:00:00	0.999 + EXPO(20.9)	-0.5 + WEIB(8.26, 1.27)
17:00:00 - 18:30:00	-0.001 + WEIB(8.72, 0.847)	-0.5 + WEIB(3.14, 1.51)
18:30:00 - 19:00:00	-0.001 + WEIB(10.8, 0.923)	-0.5 + WEIB(2.68, 1.56)
19:00:00 - 22:00:00	-0.001 + WEIB(12.2, 0.78)	EXPO(5.68)
22:00:00 - 23:59:30	-0.001 + WEIB(53.4, 0.656)	-0.001 + WEIB(37.9, 0.784)

Table 4.6 Arrivals to Konak and Ucyol on weekdays

Time period	Konak	Ucyol
05:55:00 - 06:00:00	-0.001 + WEIB(49.4, 0.678)	-0.001 + WEIB(8.08, 0.722)
06:00:00 - 07:00:00	-0.001 + WEIB(49.4, 0.678)	-0.001 + WEIB(8.08, 0.722)
07:00:00 - 07:30:00	-0.001 + WEIB(13, 0.831)	-0.5 + EXPO(2.59)
07:30:00 - 09:00:00	-0.5 + EXPO(7)	-0.5 + EXPO(1.66)
09:00:00 - 09:30:00	-0.5 + WEIB(7.15, 1.13)	-0.5 + WEIB(2.57, 1.12)
09:30:00 - 11:30:00	-0.5 + WEIB(6.1, 1.27)	-0.5 + WEIB(2.89, 1.26)
11:30:00 - 17:00:00	-0.5 + EXPO(4.13)	-0.5 + EXPO(4.37)
17:00:00 - 18:30:00	-0.5 + WEIB(3.29, 1.37)	-0.5 + WEIB(5.94, 1.13)
18:30:00 - 19:00:00	-0.5 + WEIB(3.45, 1.37)	-0.5 + WEIB(6.86, 1.17)
19:00:00 - 22:00:00	-0.5 + EXPO(6.06)	-0.001 + EXPO(13)
22:00:00 - 23:59:30	-0.001 + WEIB(13.5, 0.652)	-0.001 + EXPO(24.2)

The distributions of passenger arrivals to stations on Saturday are given in Table 4.7 - 4.11.

Table 4.7 Arrivals to Bornova and Bolge on Saturday

Time period	Bornova	Bolge
05:55:00 - 06:00:00	-0.5 + EXPO(4.53)	-0.5 + EXPO(10.6)
06:00:00 - 11:00:00	-0.5 + EXPO(4.53)	-0.5 + EXPO(10.6)
11:00:00 - 19:00:00	-0.5 + EXPO(3.14)	-0.5 + EXPO(7.85)
19:00:00 - 22:00:00	-0.5 + EXPO(8.17)	-0.001 + EXPO(20.6)
22:00:00 - 23:59:30	-0.001 + WEIB(11, 0.738)	0.999 + WEIB(38.1, 0.688)

Table 4.8 Arrivals to Sanayi and Stadyum on Saturday

Time period	Sanayi	Stadyum
05:55:00 - 06:00:00	-0.001 + EXPO(52)	-0.5 + EXPO(18)
06:00:00 - 11:00:00	-0.001 + EXPO(52)	-0.5 + EXPO(18)
11:00:00 - 19:00:00	-0.001 + EXPO(23.5)	-0.5 + EXPO(8.96)
19:00:00 - 22:00:00	0.999 + WEIB(50.1, 0.689)	-0.001 + WEIB(22.3, 0.77)
22:00:00 - 23:59:30	2 + EXPO(83.4)	-0.001 + EXPO(57.2)

Table 4.9 Arrivals to Halkapinar and Hilal on Saturday

Time period	Halkapinar	Hilal
05:55:00 - 06:00:00	-0.001 + WEIB(48.1, 0.727)	-0.001 + WEIB(60.1, 0.654)
06:00:00 - 11:00:00	-0.001 + WEIB(48.1, 0.727)	-0.001 + WEIB(60.1, 0.654)
11:00:00 - 19:00:00	-0.5 + EXPO(13.5)	-0.001 + EXPO(57.9)
19:00:00 - 22:00:00	-0.001 + WEIB(54.8, 0.67)	-0.001 + EXPO(147)
22:00:00 - 23:59:30	3 + WEIB(159, 0.329)	5 + EXPO(311)

Table 4.10 Arrivals to Basmane and Cankaya on Saturday

Time period	Basmane	Cankaya
05:55:00 - 06:00:00	-0.001 + WEIB(25, 0.775)	-0.001 + EXPO(21.5)
06:00:00 - 11:00:00	-0.001 + WEIB(25, 0.775)	-0.001 + EXPO(21.5)
11:00:00 - 19:00:00	-0.5 + EXPO(13.7)	-0.5 + EXPO(3.22)
19:00:00 - 22:00:00	-0.001 + WEIB(17.9, 0.707)	-0.5 + EXPO(8.3)
22:00:00 - 23:59:30	-0.001 + EXPO(46.2)	-0.001 + WEIB(30.7, 0.804)

Table 4.11 Arrivals to Konak and Ucyol on Saturday

Time period	Konak	Ucyol
05:55:00 - 06:00:00	-0.5 + EXPO(19.8)	-0.5 + EXPO(2.74)
06:00:00 - 11:00:00	-0.5 + EXPO(19.8)	-0.5 + EXPO(2.74)
11:00:00 - 19:00:00	-0.5 + EXPO(3.45)	-0.5 + EXPO(3.71)
19:00:00 - 22:00:00	-0.5 + EXPO(5.17)	-0.5 + EXPO(9.16)
22:00:00 - 23:59:30	-0.5 + EXPO(11.7)	-0.5 + EXPO(17.9)

The distributions of passenger arrivals to stations on Sunday are given in Table 4.12 - 4.16.

Table 4.12 Arrivals to Bornova and Bolge on Sunday

Time period	Bornova	Bolge
05:55:00 - 06:00:00	-0.5 + EXPO(11)	-0.001 + EXPO(19.9)
06:00:00 - 09:00:00	-0.5 + EXPO(11)	-0.001 + EXPO(19.9)
09:00:00 - 20:00:00	-0.5 + EXPO(4.4)	-0.5 + EXPO(13.4)
20:00:00 - 23:59:30	-0.5 + EXPO(12.9)	-0.001 + EXPO(48.2)

Table 4.13 Arrivals to Sanayi and Stadyum on Sunday

Time period	Sanayi	Stadyum
05:55:00 - 06:00:00	2 + EXPO(110)	-0.001 + EXPO(65.2)
06:00:00 - 09:00:00	2 + EXPO(110)	-0.001 + EXPO(65.2)
09:00:00 - 20:00:00	-0.001 + WEIB(44.7, 0.7)	-0.001 + EXPO(18.3)
20:00:00 - 23:59:30	0.999 + EXPO(191)	-0.001 + WEIB(56, 0.762)

Table 4.14 Arrivals to Halkapinar and Hilal on Sunday

Time period	Halkapinar	Hilal
05:55:00 - 06:00:00	2 + EXPO(214)	-0.001 + EXPO(245)
06:00:00 - 09:00:00	2 + EXPO(214)	-0.001 + EXPO(245)
09:00:00 - 20:00:00	-0.001 + EXPO(56.2)	-0.001 + EXPO(39.8)
20:00:00 - 23:59:30	0.999 + EXPO(175)	0.999 + EXPO(335)

Table 4.15 Arrivals to Basmane and Cankaya on Sunday

Time period	Basmane	Cankaya
05:55:00 - 06:00:00	0.999 + EXPO(76)	2 + EXPO(83)
06:00:00 - 09:00:00	0.999 + EXPO(76)	2 + EXPO(83)
09:00:00 - 20:00:00	-0.001 + EXPO(19.9)	-0.5 + EXPO(6.83)
20:00:00 - 23:59:30	-0.001 + EXPO(45.8)	-0.001 + EXPO(30)

Table 4.16 Arrivals to Konak and Ucyol on Sunday

Time period	Konak	Ucyol
05:55:00 - 06:00:00	-0.001 + WEIB(44.9, 0.654)	-0.001 + EXPO(15.1)
06:00:00 - 09:00:00	-0.001 + WEIB(44.9, 0.654)	-0.001 + EXPO(15.1)
09:00:00 - 20:00:00	-0.5 + EXPO(5.13)	-0.001 + EXPO(71.7)
20:00:00 - 23:59:30	-0.5 + EXPO(11.8)	-0.001 + EXPO(95)

The EXPO(mean) and WEIB(scale,shape) that are shown in Table 4.2 - 4.16 are exponential and weibull distributions respectively, and the value in front of distributions are smoothing factor.

The exponential distribution is often used to model random arrival and breakdown processes (Pedgen et al., 1990, p.561). The weibull distribution is widely used in reliability models to represent the life time of a device, is also used to represent non-negative task times that are skewed to the left (Pedgen et al., 1990, p.568).

4.2.2 Passenger Destination Probabilities

Passenger destination probability is one of the inputs of the simulation model. The probabilities based on the questionnaire conducted by the metro company with 7500 passengers between 03/12/2001 and 09/12/2001 dates are given in Table 4.17 - 4.26. For instance, in Table 4.17 and in period 06:00:00 and 07:00:00, destination probability for Bolge is 0.07. It means that a passenger boards the train at Bornova station between 06:00:00 and 07:00:00 alights the train in Bolge station with a probability of 0.07. The other probabilities are interpreted in the same manner.

Table 4.17 Destination probabilities for Bornova

Time interval	Bolge	Sanayi	Stadyum	Halkapinar	Hilal	Basmane	Cankaya	Konak	Ucyol
05:55:00 - 06:00:00	0.07	0.06	0.06	0.09	0.10	0.10	0.16	0.15	0.21
06:00:00 - 07:00:00	0.07	0.06	0.06	0.09	0.10	0.10	0.16	0.15	0.21
07:00:00 - 08:00:00	0.02	0.05	0.10	0.07	0.05	0.06	0.25	0.22	0.18
08:00:00 - 09:00:00	0.03	0.05	0.11	0.06	0.02	0.05	0.29	0.26	0.13
09:00:00 - 10:00:00	0.04	0.04	0.07	0.03	0.06	0.04	0.23	0.27	0.22
10:00:00 - 11:00:00	0.03	0.03	0.06	0.01	0.09	0.05	0.21	0.34	0.18
11:00:00 - 12:00:00	0.05	0.04	0.07	0.03	0.08	0.02	0.21	0.32	0.18
12:00:00 - 13:00:00	0.05	0.02	0.07	0.06	0.04	0.05	0.22	0.27	0.22
13:00:00 - 14:00:00	0.06	0.03	0.07	0.02	0.05	0.07	0.19	0.29	0.22
14:00:00 - 15:00:00	0.06	0.02	0.06	0.06	0.10	0.04	0.20	0.18	0.28
15:00:00 - 16:00:00	0.07	0.03	0.04	0.07	0.01	0.06	0.12	0.27	0.33
16:00:00 - 17:00:00	0.04	0.03	0.05	0.04	0.08	0.04	0.16	0.16	0.40
17:00:00 - 18:00:00	0.04	0.02	0.03	0.05	0.09	0.06	0.12	0.16	0.43
18:00:00 - 19:00:00	0.06	0.03	0.05	0.04	0.04	0.06	0.13	0.18	0.41
19:00:00 - 20:00:00	0.09	0.02	0.07	0.06	0.11	0.06	0.09	0.13	0.37
20:00:00 - 21:00:00	0.03	0.04	0.04	0.10	0.09	0.04	0.09	0.15	0.42
21:00:00 - 22:00:00	0.08	0.04	0.09	0.02	0.09	0.07	0.06	0.12	0.43
22:00:00 - 23:00:00	0.09	0.04	0.04	0.09	0.07	0.08	0.07	0.14	0.38
23:00:00 - 23:59:30	0.16	0.01	0.09	0.10	0.04	0.07	0.10	0.05	0.38

Table 4.18 Destination probabilities for Bolge

Time interval	Bornova	Sanayi	Stadyum	Halkapinar	Hilal	Basmane	Cankaya	Konak	Ucyol
05:55:00 - 06:00:00	0.26	0.04	0.04	0.07	0.08	0.08	0.13	0.12	0.18
06:00:00 - 07:00:00	0.26	0.04	0.04	0.07	0.08	0.08	0.13	0.12	0.18
07:00:00 - 08:00:00	0.32	0.03	0.07	0.05	0.04	0.04	0.17	0.16	0.12
08:00:00 - 09:00:00	0.32	0.03	0.07	0.04	0.02	0.03	0.20	0.18	0.11
09:00:00 - 10:00:00	0.26	0.03	0.05	0.03	0.05	0.03	0.18	0.21	0.16
10:00:00 - 11:00:00	0.26	0.03	0.04	0.01	0.07	0.04	0.16	0.26	0.13
11:00:00 - 12:00:00	0.27	0.03	0.05	0.02	0.07	0.02	0.16	0.25	0.13
12:00:00 - 13:00:00	0.28	0.01	0.05	0.04	0.03	0.04	0.16	0.21	0.18
13:00:00 - 14:00:00	0.27	0.02	0.05	0.02	0.04	0.05	0.15	0.23	0.17
14:00:00 - 15:00:00	0.24	0.02	0.05	0.05	0.08	0.03	0.16	0.15	0.22
15:00:00 - 16:00:00	0.23	0.02	0.03	0.06	0.01	0.05	0.10	0.23	0.27
16:00:00 - 17:00:00	0.19	0.02	0.04	0.03	0.07	0.03	0.14	0.13	0.35
17:00:00 - 18:00:00	0.23	0.02	0.02	0.04	0.07	0.04	0.09	0.13	0.36
18:00:00 - 19:00:00	0.21	0.03	0.04	0.03	0.04	0.05	0.11	0.15	0.34
19:00:00 - 20:00:00	0.25	0.02	0.06	0.05	0.09	0.05	0.08	0.10	0.30
20:00:00 - 21:00:00	0.28	0.03	0.03	0.07	0.07	0.03	0.06	0.11	0.32
21:00:00 - 22:00:00	0.25	0.04	0.08	0.02	0.08	0.06	0.05	0.10	0.32
22:00:00 - 23:00:00	0.29	0.03	0.03	0.07	0.06	0.06	0.06	0.11	0.29
23:00:00 - 23:59:30	0.29	0.01	0.07	0.08	0.04	0.06	0.08	0.04	0.33

Table 4.19 Destination probabilities for Sanayi

Time interval	Bornova	Bolge	Stadyum	Halkapinar	Hilal	Basmane	Cankaya	Konak	Ucyol
05:55:00 - 06:00:00	0.26	0.05	0.04	0.07	0.07	0.07	0.13	0.12	0.19
06:00:00 - 07:00:00	0.26	0.05	0.04	0.07	0.07	0.07	0.13	0.12	0.19
07:00:00 - 08:00:00	0.33	0.01	0.07	0.05	0.04	0.04	0.17	0.16	0.13
08:00:00 - 09:00:00	0.33	0.02	0.08	0.04	0.02	0.03	0.21	0.18	0.09
09:00:00 - 10:00:00	0.26	0.03	0.05	0.03	0.05	0.03	0.18	0.21	0.16
10:00:00 - 11:00:00	0.26	0.03	0.04	0.01	0.07	0.04	0.16	0.26	0.13
11:00:00 - 12:00:00	0.27	0.04	0.05	0.02	0.06	0.02	0.16	0.25	0.13
12:00:00 - 13:00:00	0.27	0.04	0.05	0.04	0.03	0.03	0.16	0.20	0.18
13:00:00 - 14:00:00	0.26	0.04	0.05	0.02	0.04	0.05	0.15	0.22	0.17
14:00:00 - 15:00:00	0.23	0.05	0.05	0.05	0.08	0.03	0.16	0.15	0.20
15:00:00 - 16:00:00	0.22	0.05	0.03	0.06	0.01	0.05	0.10	0.22	0.26
16:00:00 - 17:00:00	0.19	0.04	0.04	0.03	0.07	0.03	0.14	0.13	0.33
17:00:00 - 18:00:00	0.23	0.03	0.02	0.04	0.07	0.04	0.09	0.12	0.36
18:00:00 - 19:00:00	0.20	0.05	0.04	0.03	0.03	0.05	0.11	0.14	0.35
19:00:00 - 20:00:00	0.23	0.07	0.05	0.05	0.08	0.05	0.07	0.10	0.30
20:00:00 - 21:00:00	0.28	0.02	0.03	0.07	0.07	0.03	0.07	0.11	0.32
21:00:00 - 22:00:00	0.24	0.07	0.08	0.02	0.08	0.06	0.05	0.10	0.30
22:00:00 - 23:00:00	0.28	0.07	0.03	0.07	0.06	0.06	0.06	0.11	0.26
23:00:00 - 23:59:30	0.26	0.12	0.06	0.07	0.03	0.06	0.07	0.04	0.29

Table 4.20 Destination probabilities for Stadyum

Time interval	Bornova	Bolge	Sanayi	Halkapinar	Hilal	Basmane	Cankaya	Konak	Ucyol
05:55:00 - 06:00:00	0.26	0.05	0.04	0.07	0.07	0.07	0.13	0.12	0.19
06:00:00 - 07:00:00	0.26	0.05	0.04	0.07	0.07	0.07	0.13	0.12	0.19
07:00:00 - 08:00:00	0.34	0.01	0.04	0.05	0.04	0.05	0.18	0.17	0.12
08:00:00 - 09:00:00	0.34	0.02	0.03	0.04	0.02	0.03	0.22	0.19	0.11
09:00:00 - 10:00:00	0.27	0.03	0.03	0.03	0.05	0.03	0.18	0.21	0.17
10:00:00 - 11:00:00	0.26	0.03	0.03	0.01	0.07	0.04	0.16	0.26	0.14
11:00:00 - 12:00:00	0.27	0.04	0.03	0.02	0.07	0.02	0.16	0.25	0.14
12:00:00 - 13:00:00	0.29	0.04	0.01	0.04	0.03	0.04	0.17	0.21	0.17
13:00:00 - 14:00:00	0.27	0.04	0.02	0.02	0.04	0.05	0.15	0.23	0.18
14:00:00 - 15:00:00	0.24	0.05	0.02	0.05	0.08	0.03	0.16	0.15	0.22
15:00:00 - 16:00:00	0.22	0.05	0.02	0.06	0.01	0.05	0.10	0.22	0.27
16:00:00 - 17:00:00	0.19	0.04	0.02	0.03	0.07	0.03	0.14	0.14	0.34
17:00:00 - 18:00:00	0.23	0.03	0.02	0.04	0.07	0.04	0.09	0.13	0.35
18:00:00 - 19:00:00	0.21	0.05	0.02	0.03	0.04	0.05	0.11	0.15	0.34
19:00:00 - 20:00:00	0.24	0.08	0.02	0.05	0.09	0.05	0.08	0.10	0.29
20:00:00 - 21:00:00	0.28	0.02	0.03	0.07	0.07	0.03	0.07	0.11	0.32
21:00:00 - 22:00:00	0.25	0.07	0.04	0.02	0.08	0.06	0.05	0.10	0.33
22:00:00 - 23:00:00	0.28	0.07	0.03	0.07	0.06	0.06	0.06	0.11	0.26
23:00:00 - 23:59:30	0.28	0.13	0.01	0.08	0.03	0.06	0.08	0.04	0.29

Table 4.21 Destination probabilities for Halkapinar

Time interval	Bornova	Bolge	Sanayi	Stadyum	Hilal	Basmane	Cankaya	Konak	Ucyol
05:55:00 - 06:00:00	0.27	0.06	0.05	0.05	0.08	0.08	0.13	0.12	0.16
06:00:00 - 07:00:00	0.27	0.06	0.05	0.05	0.08	0.08	0.13	0.12	0.16
07:00:00 - 08:00:00	0.33	0.01	0.04	0.07	0.04	0.05	0.18	0.16	0.12
08:00:00 - 09:00:00	0.33	0.02	0.03	0.08	0.02	0.03	0.21	0.18	0.10
09:00:00 - 10:00:00	0.26	0.03	0.03	0.05	0.05	0.03	0.18	0.21	0.16
10:00:00 - 11:00:00	0.25	0.03	0.03	0.04	0.07	0.04	0.16	0.26	0.12
11:00:00 - 12:00:00	0.27	0.04	0.03	0.05	0.06	0.02	0.16	0.24	0.13
12:00:00 - 13:00:00	0.28	0.04	0.01	0.05	0.03	0.04	0.16	0.21	0.18
13:00:00 - 14:00:00	0.26	0.04	0.02	0.05	0.04	0.05	0.15	0.22	0.17
14:00:00 - 15:00:00	0.24	0.05	0.02	0.05	0.08	0.03	0.16	0.15	0.22
15:00:00 - 16:00:00	0.23	0.06	0.02	0.03	0.01	0.05	0.10	0.23	0.27
16:00:00 - 17:00:00	0.19	0.04	0.02	0.04	0.07	0.03	0.14	0.13	0.34
17:00:00 - 18:00:00	0.23	0.03	0.02	0.02	0.07	0.04	0.09	0.13	0.37
18:00:00 - 19:00:00	0.21	0.05	0.02	0.04	0.04	0.05	0.11	0.15	0.33
19:00:00 - 20:00:00	0.24	0.08	0.02	0.05	0.09	0.05	0.08	0.10	0.29
20:00:00 - 21:00:00	0.29	0.02	0.03	0.03	0.07	0.03	0.07	0.12	0.34
21:00:00 - 22:00:00	0.24	0.07	0.03	0.07	0.07	0.05	0.05	0.10	0.32
22:00:00 - 23:00:00	0.29	0.07	0.03	0.03	0.06	0.06	0.06	0.11	0.29
23:00:00 - 23:59:30	0.28	0.13	0.01	0.07	0.03	0.06	0.08	0.04	0.30

Table 4.22 Destination probabilities for Hilal

Time interval	Bornova	Bolge	Sanayi	Stadyum	Halkapinar	Basmane	Cankaya	Konak	Ucyol
05:55:00 - 06:00:00	0.27	0.06	0.05	0.05	0.07	0.08	0.13	0.12	0.17
06:00:00 - 07:00:00	0.27	0.06	0.05	0.05	0.07	0.08	0.13	0.12	0.17
07:00:00 - 08:00:00	0.33	0.01	0.04	0.07	0.05	0.04	0.18	0.16	0.12
08:00:00 - 09:00:00	0.32	0.02	0.03	0.07	0.04	0.03	0.20	0.18	0.11
09:00:00 - 10:00:00	0.27	0.03	0.03	0.05	0.03	0.03	0.18	0.21	0.17
10:00:00 - 11:00:00	0.27	0.03	0.03	0.05	0.01	0.04	0.17	0.27	0.13
11:00:00 - 12:00:00	0.28	0.04	0.03	0.05	0.02	0.02	0.17	0.25	0.14
12:00:00 - 13:00:00	0.28	0.04	0.01	0.05	0.04	0.04	0.16	0.20	0.18
13:00:00 - 14:00:00	0.26	0.04	0.02	0.05	0.02	0.05	0.15	0.22	0.19
14:00:00 - 15:00:00	0.24	0.05	0.02	0.05	0.05	0.03	0.17	0.15	0.24
15:00:00 - 16:00:00	0.22	0.05	0.02	0.03	0.06	0.05	0.09	0.22	0.26
16:00:00 - 17:00:00	0.20	0.04	0.02	0.05	0.03	0.03	0.14	0.14	0.35
17:00:00 - 18:00:00	0.24	0.03	0.02	0.02	0.04	0.05	0.10	0.13	0.37
18:00:00 - 19:00:00	0.21	0.05	0.02	0.04	0.03	0.05	0.11	0.15	0.34
19:00:00 - 20:00:00	0.25	0.08	0.02	0.06	0.05	0.05	0.08	0.11	0.30
20:00:00 - 21:00:00	0.29	0.02	0.03	0.03	0.08	0.03	0.07	0.12	0.33
21:00:00 - 22:00:00	0.25	0.07	0.04	0.08	0.02	0.06	0.05	0.10	0.33
22:00:00 - 23:00:00	0.29	0.07	0.03	0.03	0.07	0.06	0.06	0.11	0.28
23:00:00 - 23:59:30	0.27	0.12	0.01	0.07	0.07	0.06	0.07	0.04	0.29

Table 4.23 Destination probabilities for Basmane

Time interval	Bornova	Bolge	Sanayi	Stadyum	Halkapinar	Hilal	Cankaya	Konak	Ucyol
05:55:00 - 06:00:00	0.27	0.06	0.05	0.05	0.07	0.08	0.13	0.12	0.17
06:00:00 - 07:00:00	0.27	0.06	0.05	0.05	0.07	0.08	0.13	0.12	0.17
07:00:00 - 08:00:00	0.33	0.01	0.04	0.07	0.05	0.04	0.18	0.16	0.12
08:00:00 - 09:00:00	0.33	0.02	0.03	0.08	0.04	0.02	0.21	0.18	0.09
09:00:00 - 10:00:00	0.26	0.03	0.03	0.05	0.03	0.05	0.18	0.21	0.16
10:00:00 - 11:00:00	0.26	0.03	0.03	0.04	0.01	0.07	0.16	0.26	0.14
11:00:00 - 12:00:00	0.27	0.04	0.03	0.05	0.02	0.06	0.16	0.24	0.13
12:00:00 - 13:00:00	0.28	0.04	0.01	0.05	0.04	0.03	0.16	0.20	0.19
13:00:00 - 14:00:00	0.27	0.04	0.02	0.05	0.02	0.04	0.15	0.23	0.18
14:00:00 - 15:00:00	0.23	0.05	0.02	0.05	0.05	0.08	0.16	0.15	0.21
15:00:00 - 16:00:00	0.22	0.06	0.02	0.03	0.06	0.01	0.10	0.22	0.28
16:00:00 - 17:00:00	0.19	0.04	0.02	0.04	0.03	0.07	0.14	0.13	0.34
17:00:00 - 18:00:00	0.23	0.03	0.02	0.02	0.04	0.07	0.09	0.13	0.37
18:00:00 - 19:00:00	0.21	0.05	0.03	0.04	0.03	0.04	0.11	0.15	0.34
19:00:00 - 20:00:00	0.24	0.08	0.02	0.05	0.05	0.09	0.08	0.10	0.29
20:00:00 - 21:00:00	0.28	0.02	0.03	0.03	0.07	0.07	0.07	0.11	0.32
21:00:00 - 22:00:00	0.25	0.07	0.04	0.08	0.02	0.08	0.05	0.10	0.31
22:00:00 - 23:00:00	0.29	0.07	0.03	0.03	0.07	0.06	0.06	0.11	0.28
23:00:00 - 23:59:30	0.27	0.12	0.01	0.07	0.08	0.03	0.08	0.04	0.30

Table 4.24 Destination probabilities for Cankaya

Time interval	Bornova	Bolge	Sanayi	Stadyum	Halkapinar	Hilal	Basmane	Konak	Ucyol
05:55:00 - 06:00:00	0.28	0.06	0.05	0.05	0.07	0.08	0.08	0.13	0.20
06:00:00 - 07:00:00	0.28	0.06	0.05	0.05	0.07	0.08	0.08	0.13	0.20
07:00:00 - 08:00:00	0.38	0.02	0.04	0.08	0.06	0.04	0.05	0.19	0.14
08:00:00 - 09:00:00	0.40	0.02	0.04	0.09	0.05	0.02	0.04	0.22	0.12
09:00:00 - 10:00:00	0.31	0.03	0.03	0.06	0.03	0.05	0.03	0.24	0.22
10:00:00 - 11:00:00	0.30	0.03	0.03	0.05	0.01	0.08	0.05	0.30	0.15
11:00:00 - 12:00:00	0.31	0.05	0.04	0.06	0.02	0.07	0.02	0.28	0.15
12:00:00 - 13:00:00	0.32	0.04	0.02	0.06	0.05	0.03	0.04	0.23	0.21
13:00:00 - 14:00:00	0.30	0.05	0.02	0.06	0.02	0.04	0.06	0.25	0.20
14:00:00 - 15:00:00	0.27	0.06	0.02	0.06	0.06	0.09	0.04	0.17	0.23
15:00:00 - 16:00:00	0.24	0.06	0.02	0.03	0.06	0.01	0.05	0.24	0.29
16:00:00 - 17:00:00	0.21	0.04	0.03	0.05	0.04	0.08	0.04	0.15	0.36
17:00:00 - 18:00:00	0.25	0.03	0.02	0.02	0.04	0.08	0.05	0.13	0.38
18:00:00 - 19:00:00	0.22	0.05	0.03	0.04	0.04	0.04	0.05	0.16	0.37
19:00:00 - 20:00:00	0.25	0.08	0.02	0.06	0.05	0.09	0.05	0.10	0.30
20:00:00 - 21:00:00	0.29	0.02	0.03	0.03	0.08	0.07	0.03	0.12	0.33
21:00:00 - 22:00:00	0.24	0.07	0.04	0.08	0.02	0.08	0.06	0.10	0.31
22:00:00 - 23:00:00	0.29	0.07	0.03	0.03	0.07	0.06	0.06	0.11	0.28
23:00:00 - 23:59:30	0.28	0.13	0.01	0.07	0.08	0.03	0.06	0.04	0.30

Table 4.25 Destination probabilities for Konak

Time interval	Bornova	Bolge	Sanayi	Stadyum	Halkapinar	Hilal	Basmane	Cankaya	Ucyol
05:55:00 - 06:00:00	0.28	0.06	0.05	0.05	0.07	0.08	0.08	0.14	0.19
06:00:00 - 07:00:00	0.28	0.06	0.05	0.05	0.07	0.08	0.08	0.14	0.19
07:00:00 - 08:00:00	0.37	0.02	0.04	0.08	0.06	0.04	0.05	0.20	0.14
08:00:00 - 09:00:00	0.39	0.02	0.04	0.09	0.05	0.02	0.04	0.24	0.11
09:00:00 - 10:00:00	0.32	0.03	0.03	0.06	0.03	0.06	0.03	0.22	0.22
10:00:00 - 11:00:00	0.34	0.03	0.03	0.06	0.01	0.09	0.05	0.21	0.18
11:00:00 - 12:00:00	0.34	0.05	0.04	0.06	0.03	0.08	0.02	0.21	0.17
12:00:00 - 13:00:00	0.34	0.05	0.02	0.07	0.05	0.04	0.04	0.20	0.19
13:00:00 - 14:00:00	0.33	0.05	0.03	0.06	0.02	0.05	0.06	0.18	0.22
14:00:00 - 15:00:00	0.26	0.06	0.02	0.06	0.06	0.09	0.04	0.18	0.23
15:00:00 - 16:00:00	0.27	0.07	0.03	0.04	0.07	0.01	0.06	0.12	0.33
16:00:00 - 17:00:00	0.21	0.04	0.03	0.05	0.04	0.08	0.04	0.15	0.36
17:00:00 - 18:00:00	0.25	0.03	0.02	0.03	0.04	0.08	0.05	0.10	0.40
18:00:00 - 19:00:00	0.23	0.06	0.03	0.04	0.04	0.04	0.06	0.12	0.38
19:00:00 - 20:00:00	0.25	0.08	0.02	0.06	0.05	0.09	0.05	0.08	0.32
20:00:00 - 21:00:00	0.31	0.03	0.03	0.03	0.08	0.08	0.03	0.07	0.34
21:00:00 - 22:00:00	0.26	0.07	0.04	0.08	0.02	0.08	0.06	0.05	0.34
22:00:00 - 23:00:00	0.30	0.07	0.03	0.03	0.07	0.06	0.06	0.06	0.32
23:00:00 - 23:59:30	0.27	0.12	0.01	0.07	0.07	0.03	0.06	0.07	0.30

Table 4.26 Destination probabilities for Ucyol

Time interval	Bornova	Bolge	Sanayi	Stadyum	Halkapinar	Hilal	Basmane	Cankaya	Konak
05:55:00 - 06:00:00	0.30	0.06	0.05	0.05	0.08	0.09	0.09	0.15	0.13
06:00:00 - 07:00:00	0.30	0.06	0.05	0.05	0.08	0.09	0.09	0.15	0.13
07:00:00 - 08:00:00	0.35	0.02	0.04	0.08	0.06	0.04	0.05	0.19	0.17
08:00:00 - 09:00:00	0.34	0.02	0.04	0.08	0.04	0.02	0.04	0.22	0.20
09:00:00 - 10:00:00	0.32	0.03	0.03	0.06	0.03	0.05	0.03	0.21	0.24
10:00:00 - 11:00:00	0.28	0.03	0.03	0.05	0.01	0.08	0.05	0.18	0.29
11:00:00 - 12:00:00	0.29	0.05	0.04	0.06	0.02	0.07	0.02	0.18	0.27
12:00:00 - 13:00:00	0.33	0.04	0.02	0.06	0.05	0.03	0.04	0.19	0.24
13:00:00 - 14:00:00	0.32	0.05	0.02	0.06	0.02	0.04	0.06	0.17	0.26
14:00:00 - 15:00:00	0.28	0.06	0.02	0.06	0.06	0.10	0.04	0.20	0.18
15:00:00 - 16:00:00	0.28	0.07	0.03	0.04	0.08	0.01	0.07	0.13	0.29
16:00:00 - 17:00:00	0.27	0.05	0.03	0.06	0.05	0.10	0.05	0.20	0.19
17:00:00 - 18:00:00	0.34	0.05	0.02	0.03	0.05	0.11	0.07	0.14	0.19
18:00:00 - 19:00:00	0.30	0.07	0.04	0.05	0.05	0.05	0.07	0.16	0.21
19:00:00 - 20:00:00	0.32	0.10	0.02	0.07	0.07	0.11	0.07	0.10	0.14
20:00:00 - 21:00:00	0.40	0.03	0.04	0.04	0.10	0.10	0.04	0.09	0.16
21:00:00 - 22:00:00	0.35	0.09	0.05	0.10	0.02	0.10	0.08	0.07	0.14
22:00:00 - 23:00:00	0.38	0.09	0.04	0.04	0.09	0.07	0.08	0.07	0.14
23:00:00 - 23:59:30	0.37	0.16	0.01	0.09	0.10	0.04	0.08	0.10	0.05

4.2.3 Failure and Repair Time Distributions

The trains may change the current track by using switches for some reasons. For instance, there may be a failure in a part of track (block) between two sequential stations, or another train that has opposite direction may running on the next part of the current track.

If a train changes its current track, the new track will be closed to trains in opposite direction. Input Analyser module of Arena V2.2 is also used for chi-squared goodness of fit tests ($\alpha = 0.05$) for failure time between two consecutive failures in blocks, and repair times.

Table 4.27 shows the distributions of failure times and repair times according to blocks.

Table 4.27 Distributions of failure time and repair time

Block name	Probability of failure	Probability of repair time
BOL11	1.3e+005 + WEIB(3.26e+006, 0.569)	120 + WEIB(612, 0.517)
SAN11	1.75e+006 + WEIB(3.98e+006, 0.356)	180 + EXPO(2.11e+003)
STA11	2.59e+005 + WEIB(1.98e+006, 0.259)	120 + EXPO(960)
HAL13	5.1e+003 + WEIB(2.06e+006, 0.586)	60 + EXPO(1.2e+003)
HIL11	2.6e+004 + WEIB(3.05e+006, 0.443)	120 + EXPO(1.59e+003)
CAN11	1.36e+006 + WEIB(4.21e+006, 0.409)	120 + EXPO(852)
KON11	4.54e+005 + WEIB(1.99e+004, 0.136)	420 + EXPO(662)
UCY11	1.3e+005 + WEIB(2.2e+006, 0.449)	120 + EXPO(931)
UCY24	6.48e+004 + WEIB(2.2e+006, 0.666)	60 + EXPO(1.3e+003)
KON22	1.94e+005 + WEIB(2.98e+006, 0.484)	120 + EXPO(1.15e+003)
CAN22	3.05e+006 + WEIB(5.22e+006, 0.297)	300 + EXPO(523)
BAS22	1.94e+005 + WEIB(3.28e+006, 0.296)	180 + EXPO(1.7e+003)
HIL22	1.07e+004 + WEIB(3.51e+006, 0.466)	180 + WEIB(1.44e+003, 0.495)
HAL24	3.05e+006 + WEIB(4.2e+006, 0.287)	180 + EXPO(2.39e+003)
HAL42	6.48e+004 + WEIB(2.85e+006, 0.377)	180 + EXPO(3.27e+003)
SAN22	1.04e+006 + WEIB(3.71e+006, 0.365)	120 + EXPO(1.37e+003)
BOL22	1.3e+006 + WEIB(4.4e+006, 0.212)	120 + EXPO(636)

4.3 Assumptions

- The first passenger is allowed to enter the system 5 minutes before the first train's departure time. Also, last passenger is allowed to leave the system 0.5 minutes before the last train's departure time.

- Arrival time distributions and destination station probabilities for passengers who get into the system between 05:55:00 and 06:00:00 are same as first time period's values.

- The train, which leaves system to cause a decrease in the number of running trains, unloads its passengers at Halkapinar station. While a part of the unloaded passengers wait a new train to go on their trip, another part may leave the system without any delay if Halkapinar station is their destination.

- Train speed can be different at different parts of tracks, but it does not change during a trip between two successive stations. The train speeds between consecutive stations are given in Table 4.28.

Table 4.28 Train speeds between stations

Stations		Speed	
Departure	**Destination**	**(km/h)**	**(m/sec)**
End point	Bornova	80	22.22
Bornova	Bolge	80	22.22
Bolge	Sanayi	80	22.22
Sanayi	Stadyum	80	22.22
Stadyum	Halkapinar	80	22.22
Halkapinar	Hilal	80	22.22
Hilal	Basmane	80	22.22
Basmane	Cankaya	80	22.22
Cankaya	Konak	80	22.22
Konak	Ucyol	80	22.22
Ucyol	End point	40	11.11
End point	Ucyol	40	11.11
Ucyol	Konak	60	16.67
Konak	Cankaya	80	22.22
Cankaya	Basmane	80	22.22
Basmane	Hilal	80	22.22
Hilal	Halkapinar	80	22.22
Halkapinar	Stadyum	80	22.22
Stadyum	Sanayi	80	22.22
Sanayi	Bolge	80	22.22
Bolge	Bornova	80	22.22
Bornova	End point	20	5.56

- Train dwell times are static and are given in Table 3.4.

- Only one train can locate on a platform. Therefore, except Halkapinar station, which has three platforms, there can be only two trains in a station at the same time. At Halkapinar station, there can be only three trains at the same time.

- All the trains have 3 carriages, and a carriage includes 44 seats and 36 square meter empty spaces where passengers can stand. The desirable number of passengers per square meter is 4. Base on that information, we assume that a carriage can take maximum (44+36*4=188) 188 passengers at a time. In other words, the full capacity of each carriage is 188 passengers, and the full capacity of a train is 564 passengers.

- The decision on which blocks and switches will be used at the next trip is made during dwell time, and it does not change until the next station.

- If there is more than one train at the end station, the train, which comes first, leaves first.

- At stations, queue discipline is FIFO (first in first out). If the train is full and some passengers can not board the train, they have priority for the next train.

- In the simulation model, the metrics used for length is meter, and for time is second.

- Because of some extraordinary causes, in one of the end stations if there is no train to board at departure time, passengers go on waiting for a train, and when a train arrives, passengers get on and trip begins without losing any time.

4.4 Flowcharts for Events

The followings are the six main events;

- Passenger's arrival to the station
- Train's arrival to the station
- Train's departure time
- Alighting the train
- Boarding the train
- Train's departure from the station

There are also other events like System opening and System closing occurring only once during a day. Figures 4.1 - 4.5 display the flow diagrams of the main events.

Four types of entities flow between Arena Blocks. First type represents passenger, and is created until the system is closed for passenger arrivals. The next type captures a train when needed and does not release until the train leaves the system. Type three is used for calculating the number of trains needed for the first time period, moving the trains to scheduled stations before the first time period begins, and also calculating the fullness rate of the carriages. In addition, it is used for adjusting the

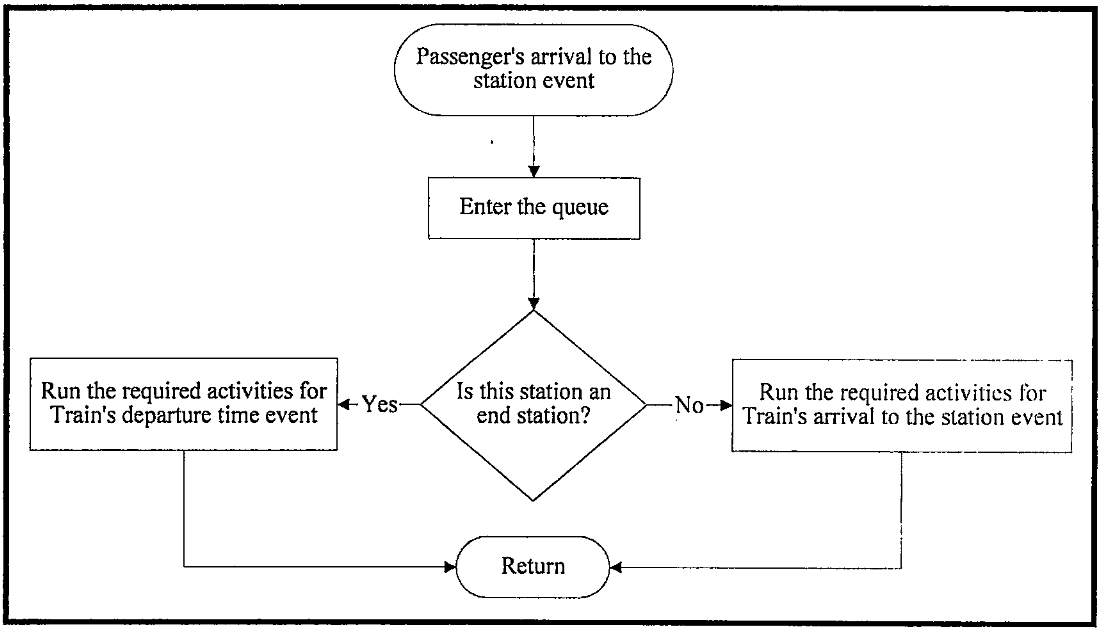

Figure 4.1 Flowchart for Passenger's arrival to the station event

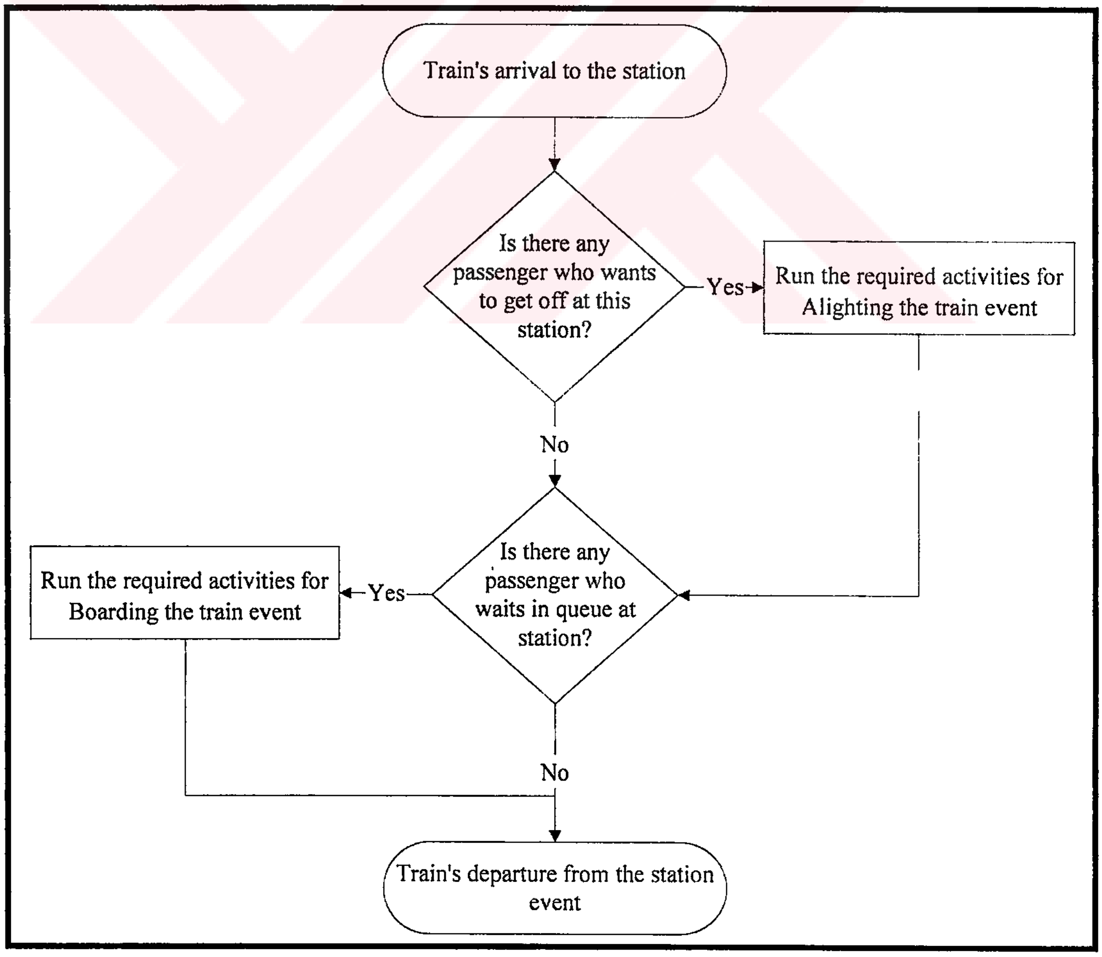

Figure 4.2 Flowchart for Train's arrival to the station event and Train's departure from the station event

number of trains at transition between two consecutive time periods, determining the departure times due to headways, and ordering the trains to departure. Fourth type is used for failures that occur at blocks, it prevents the use of failured parts, and causes a delay during repair action.

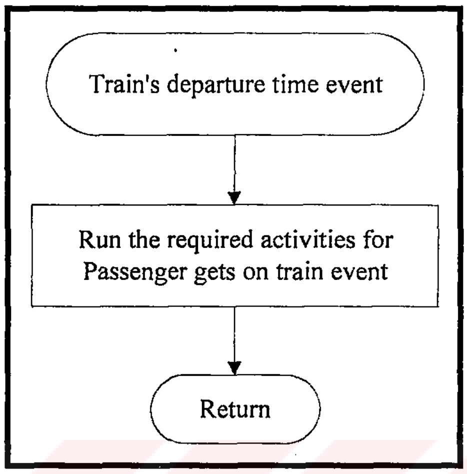

Figure 4.3 Flowchart for Train's departure time event

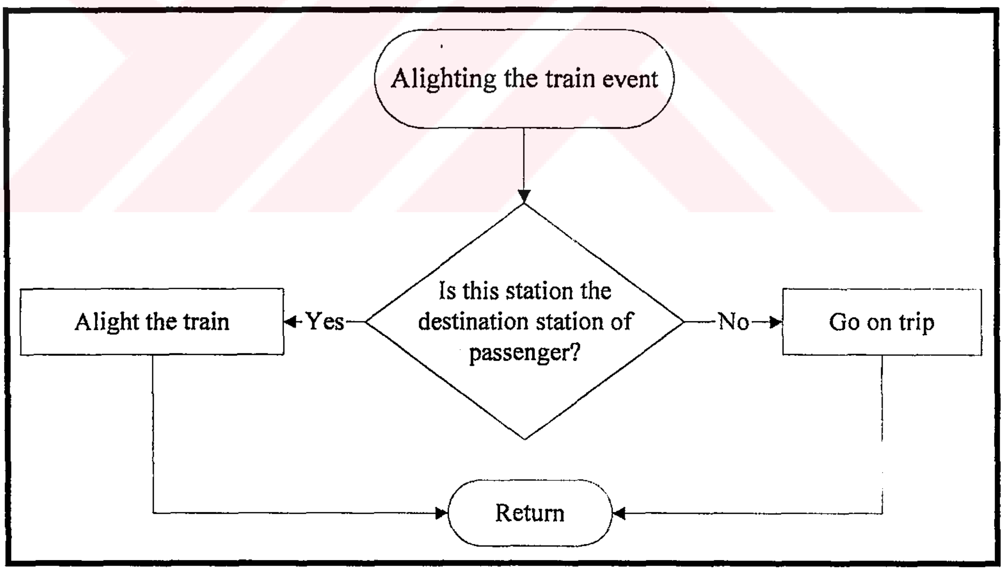

Figure 4.4 Flowchart for Alighting the train event

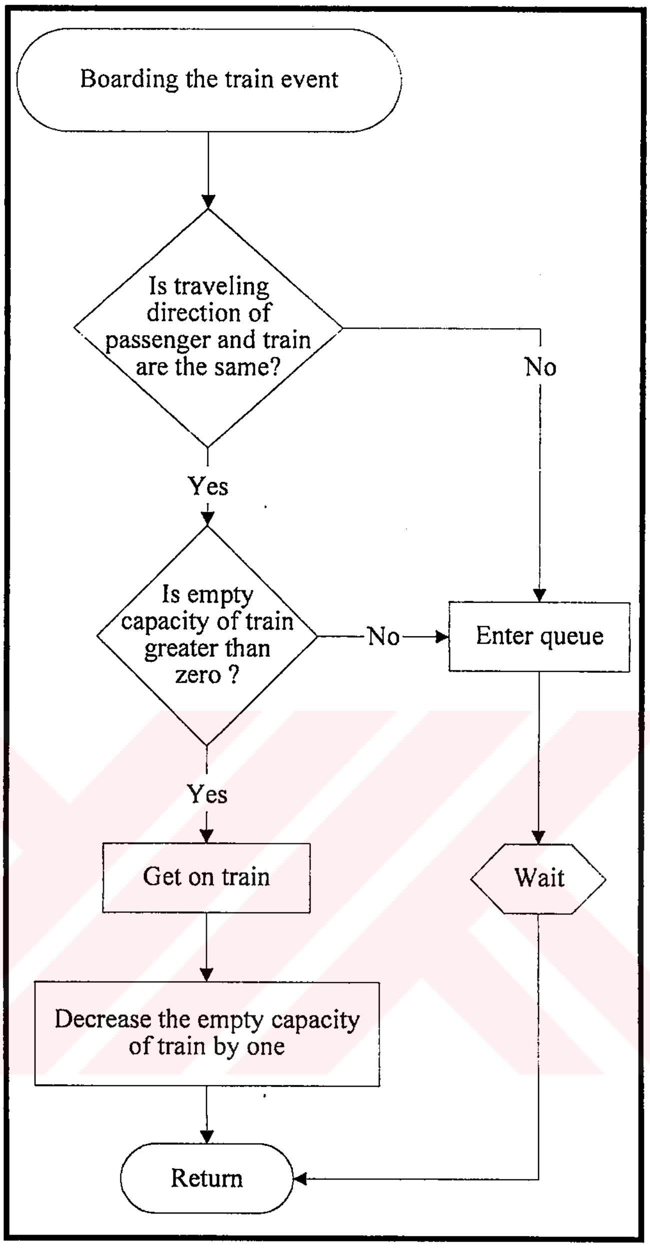

Figure 4.5 Flowchart for Boarding the train event

4.5 Flowcharts Related to Halkapinar Station

For simulation modelling, the number of Arena Blocks used is more than 3000. Only a part of the flow diagram related to Halkapinar station is given in Figures 4.6 - 4.13 for the sake of simplicity. These figures show the train scheduling procedure at Halkapinar station. This station is the most complex one. It behaves like an end station at the beginning of the first period, and also like a middle station during a day. In addition, trains use this station when they enter or leave the system. Also, there exist three platforms in Halkapinar station. In flowcharts platform number is denoted with p.

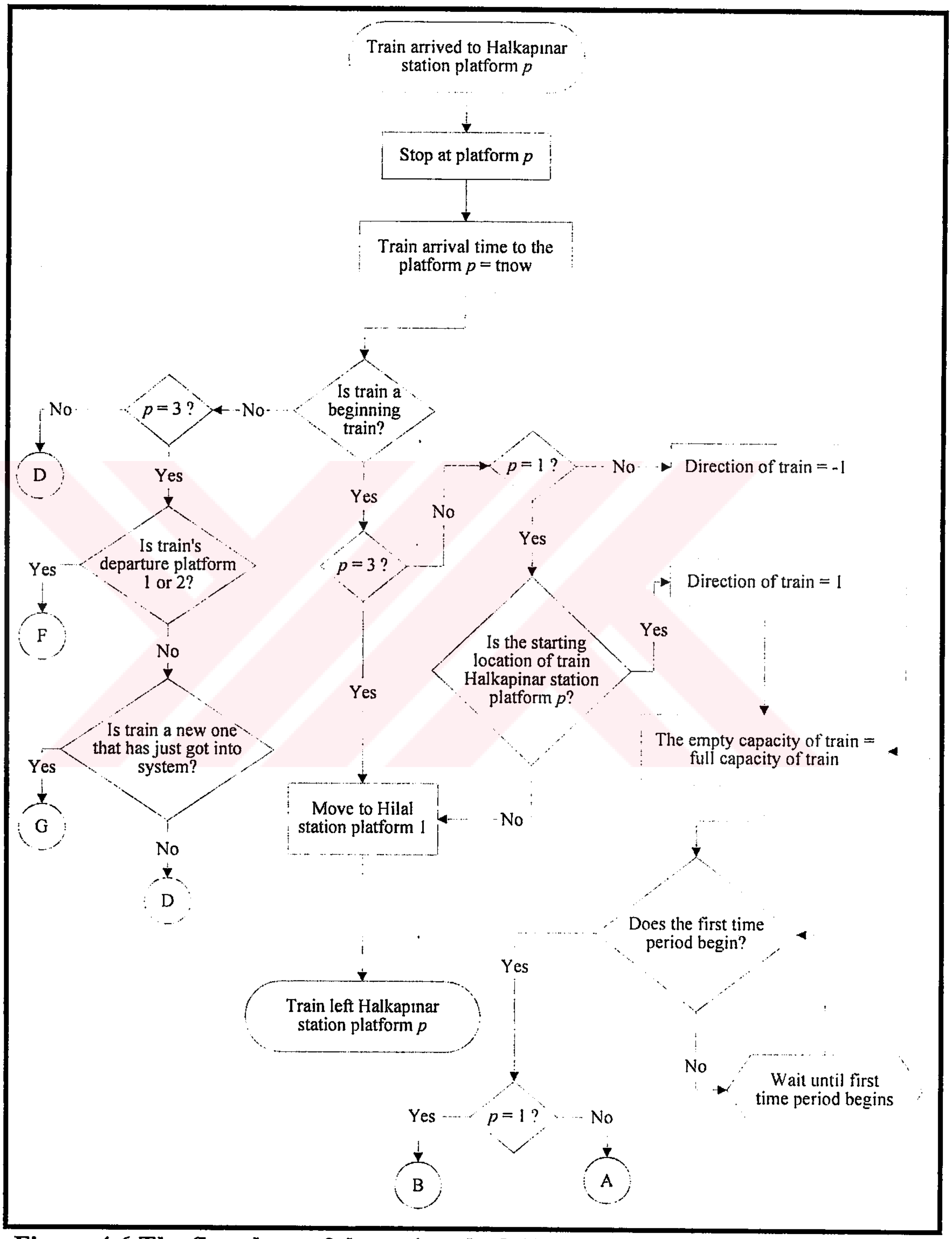

Figure 4.6 The flowchart of the train scheduling procedure (*tsp*) in Halkapinar station (part 1) (p=1,2,3)

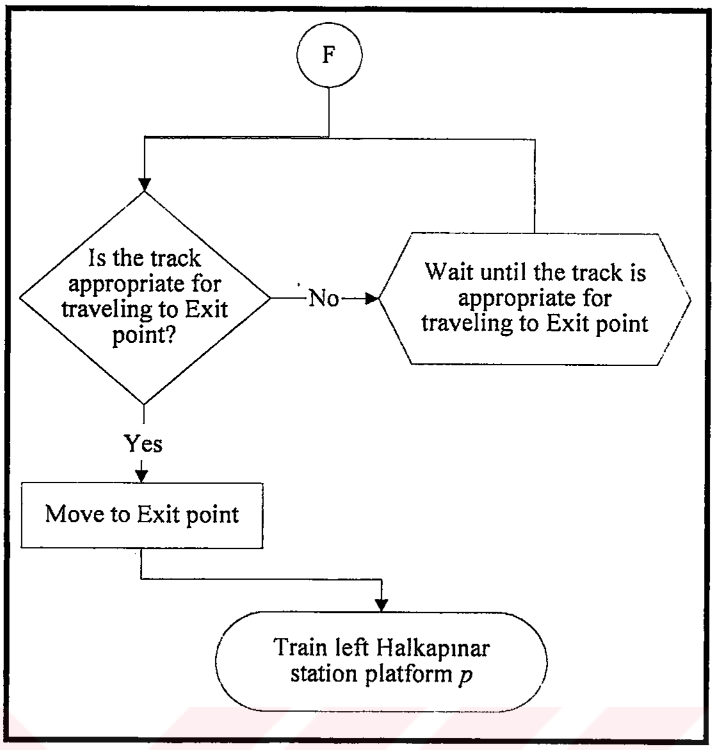

Figure 4.7 The flowchart of the *tsp* in Halkapinar station (part 2) (*p*=1,2,3)

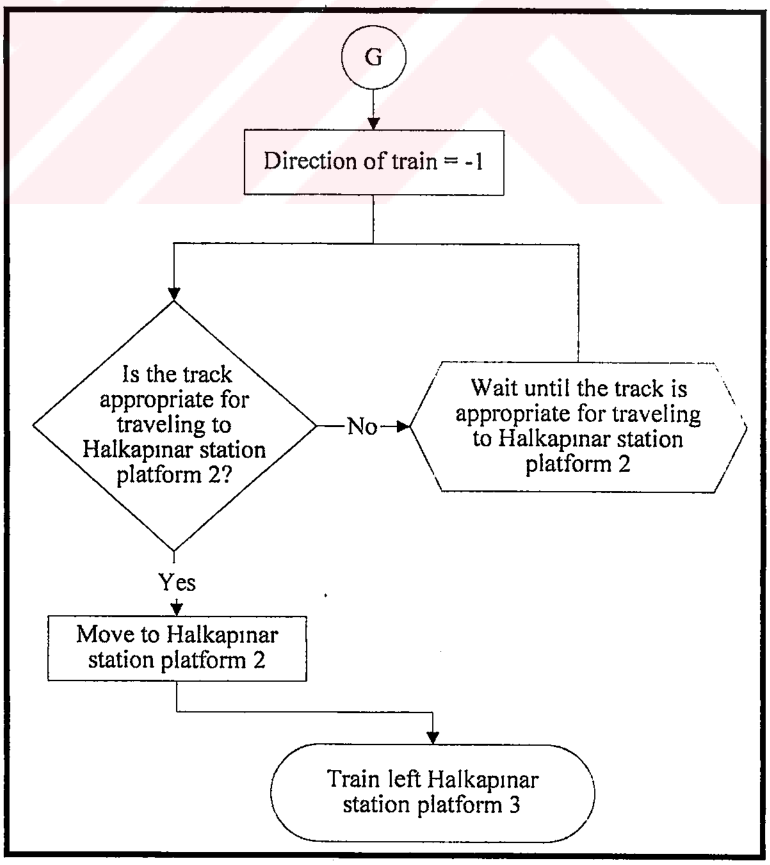

Figure 4.8 The flowchart of the *tsp* in Halkapinar station (part 3) (*p*=1,2,3)

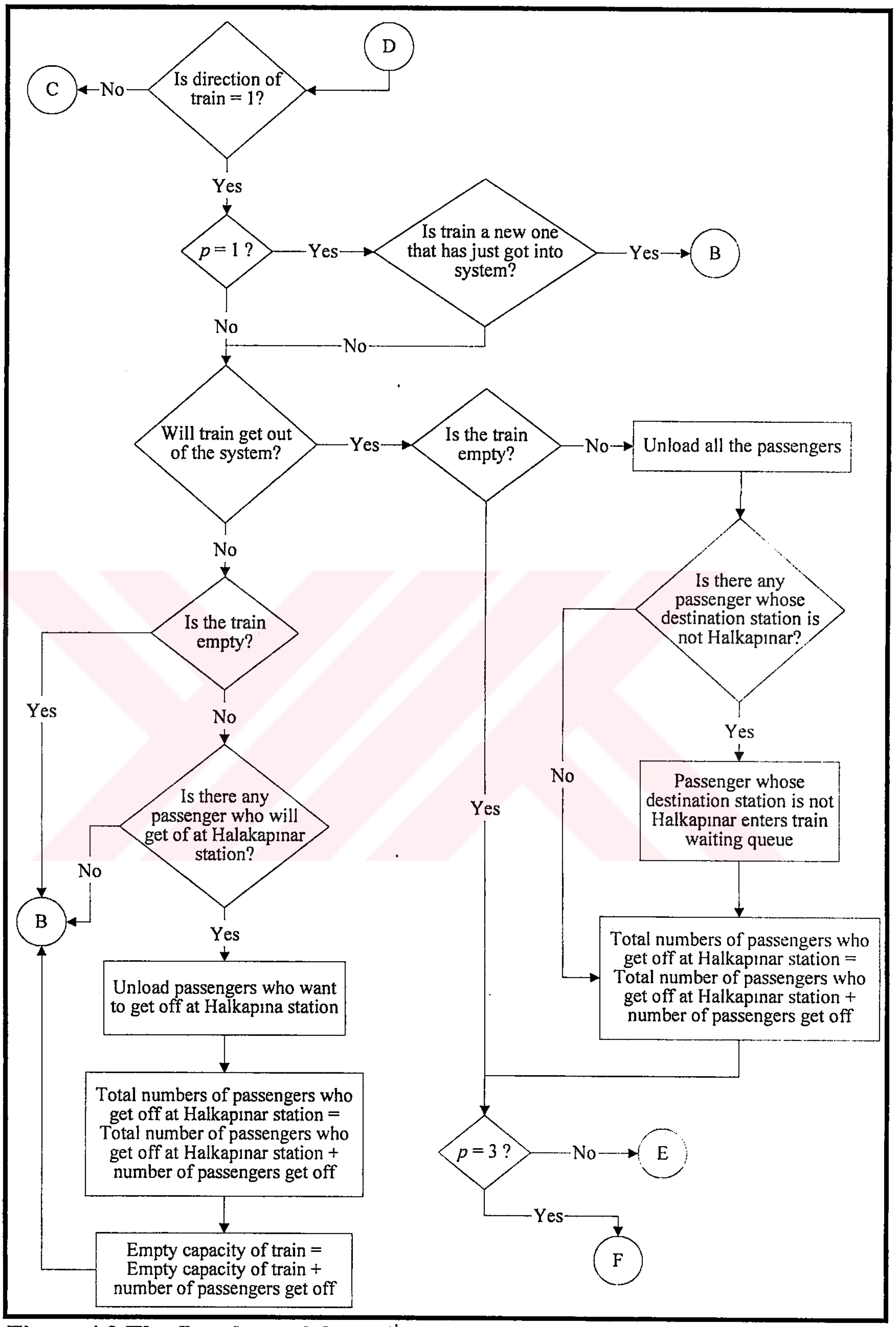

Figure 4.9 The flowchart of the train *tsp* in Halkapinar station (part 4) (*p*=1,2,3)

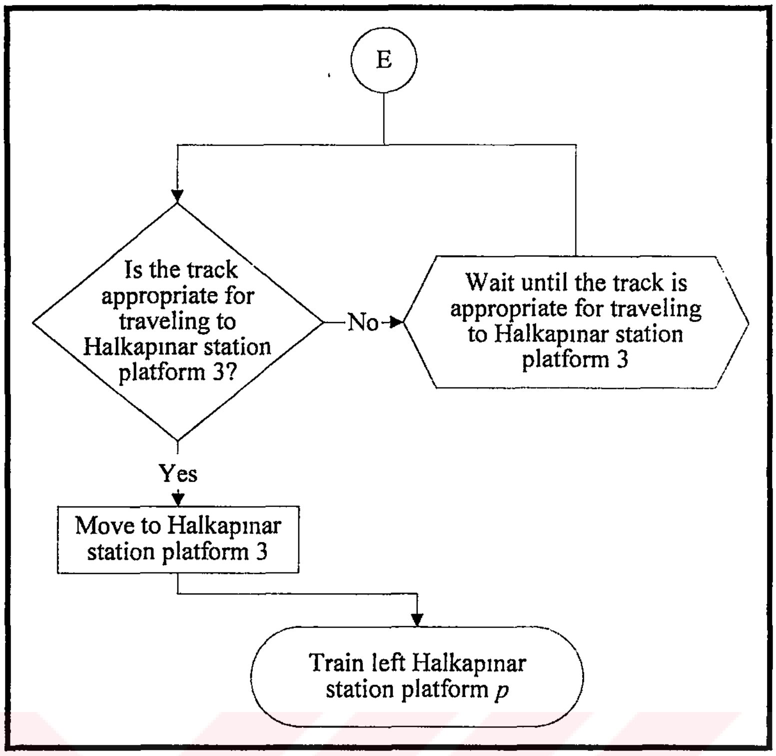

Figure 4.10 The flowchart of the *tsp* in Halkapinar station (part 5) (*p*=1,2,3)

As it can be seen from flowcharts, train which uses Halkapinar station may be a beginning train, a new one that has just entered the system, a running train that will exit the system, or a train running in Bornova or Ucyol direction.

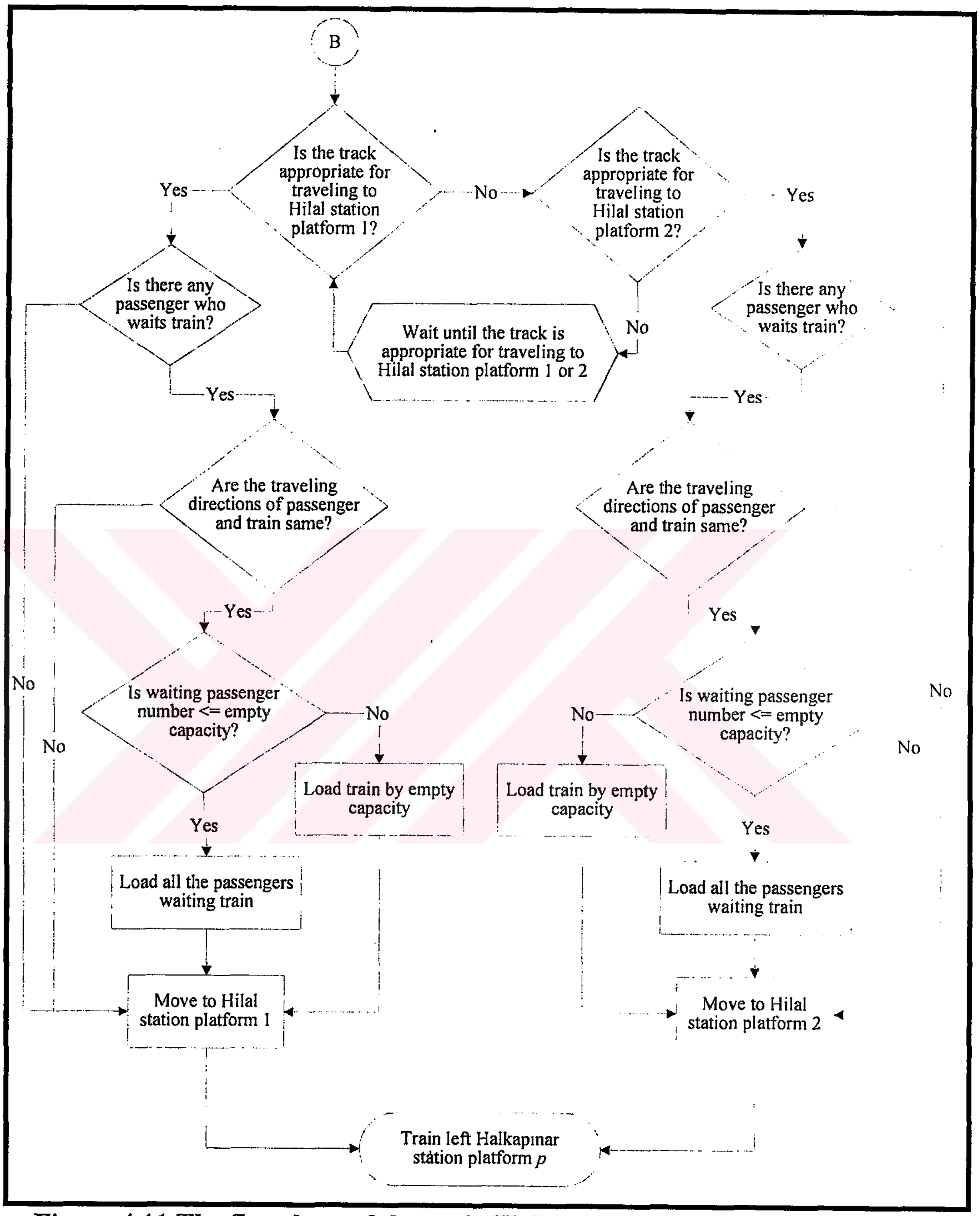

Figure 4.11 The flowchart of the *tsp* in Halkapinar station (part 6) (*p*=1,2,3)

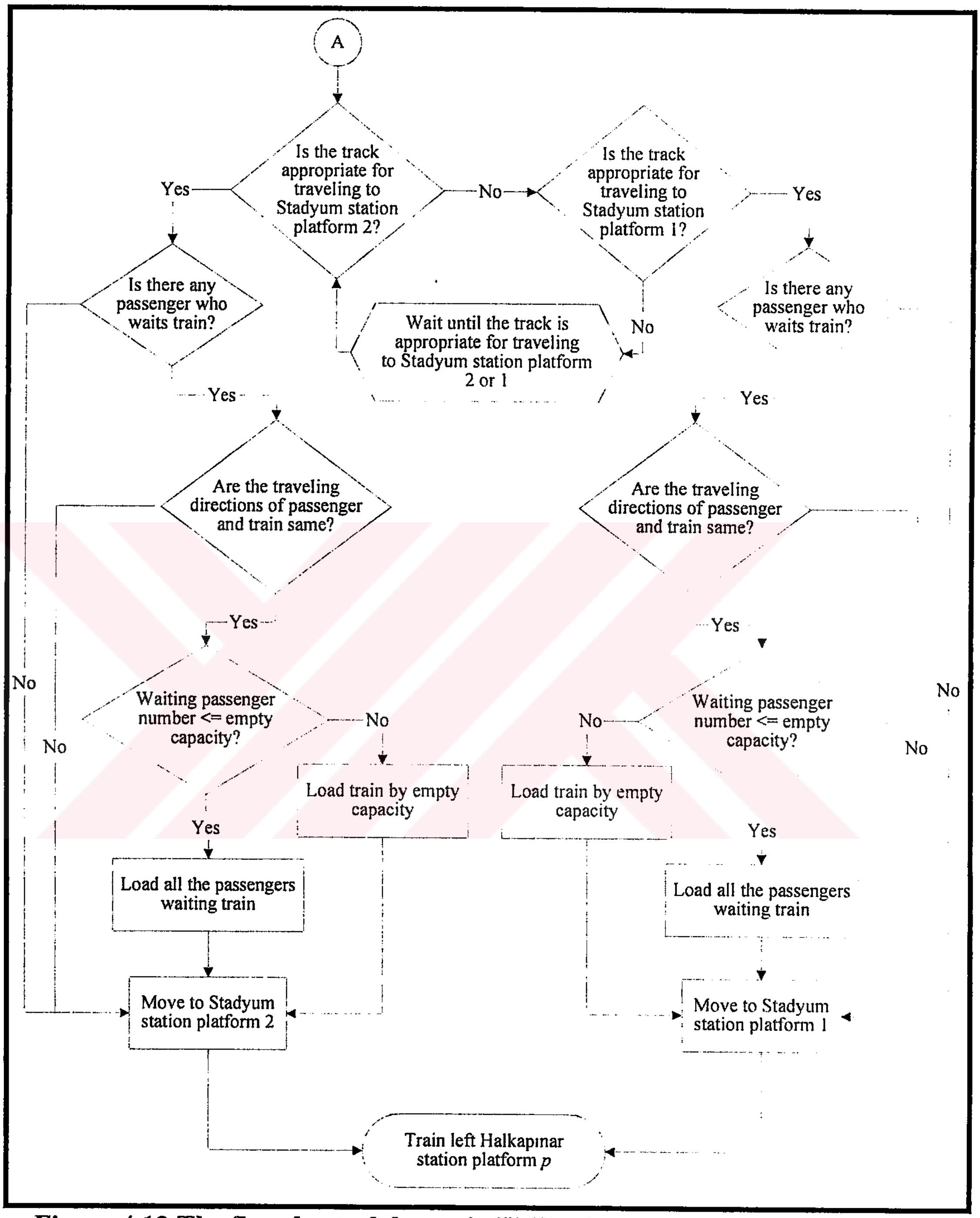

Figure 4.12 The flowchart of the *tsp* in Halkapinar station (part 7) (p=1,2,3)

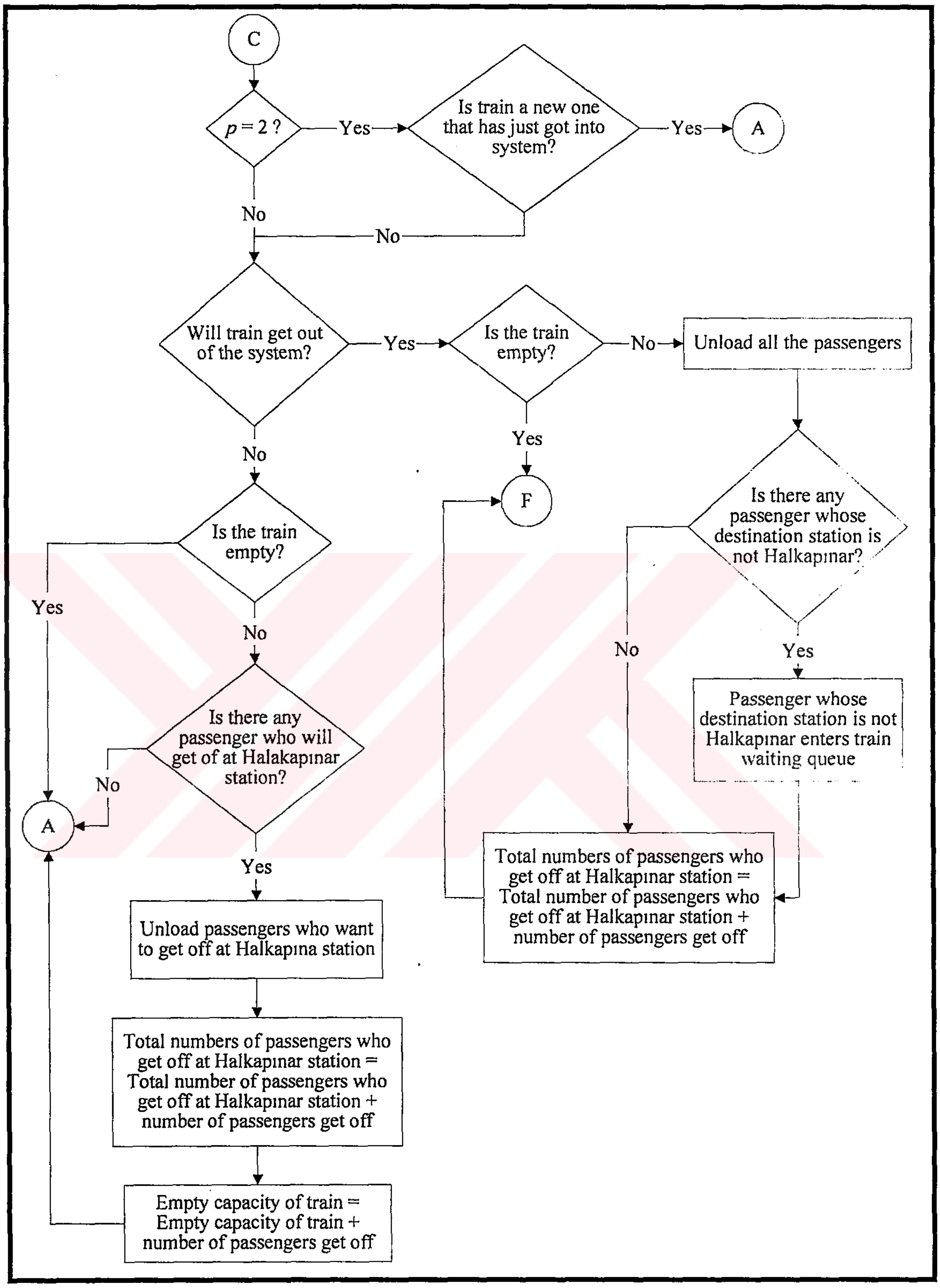

Figure 4.13 The flowchart of the *tsp* in Halkapinar station (part 8) (*p*=1,2,3)

4.6 Attributes and Variables

The attributes and variables that are used in simulation model are given below.

4.6.1 Attributes

Trigger: If an entity is a trigger entity its value for this attribute is 1, and for other entities its value is 0.

Timein: All the entities that represent passenger have this attribute. This attribute indicates the created time of an entity.

Dep_station: The value of this attribute indicates the station number that train comes from.

Empty_capacity: Indicates the empty capacity of a train. Its maximum value can be 564.

Train_close: This attribute takes values 0 and 1. The 1 indicates that train will depart the system from Halkapinar station because the system has closed. Otherwise, the value of attribute is 0.

Train_begin_no: This attribute ranges from 0 to 8. While 0 indicates that the train is not a beginning train, the values from 1 to 8 indicate that train is a beginning train. According to these values, trains' locations are determined before the first time period begins.

Direction_travel: This attribute has two values; -1 and 1. The value 1 indicates that train run towards the Ucyol station, and the value -1 indicates that train is in Bornova direction.

Des_station: Shows the station number that passenger travels to.

Train#: There can be maximum 8 trains in the system, because the minimum headway is 5 minutes. Therefore, this attribute denoting the train number ranges from 1 to 8.

4.6.2 Variables

The variables which control the blocks are shown in Table 4.29.

Table 4.29 Variables that control the blocks

Variable	Controlled block	Value interval
uba01	UCY03	[-8,8]
uba02	UCY51, UCY01	[-8,8]
uba03	UCY11, UCY13, KON01	[-8,8]
uba04	KON51, KON11, CAN01, CAN11,	[-8,8]
uba05	BAS51, BAS11, HIL01, HIL11	[-8,8]
uba06	HAL51, HAL01	[-8,8]
uba07	HAL55, HAL13	[-8,8]
uba08	HAL59, HAL41, STA01, STA11	[-8,8]
uba09	SAN51, SAN01	[-8,8]
uba10	SAN11, BOL01, BOL11	[-8,8]
uba11	BOR51, BOR01	[-8,8]
uba12	BOR03	[-8,8]
bua01	BOR04	[-8,8]
bua02	BOR02	[-8,8]
bua03	BOR52, BOL22, BOL02, SAN22,	[-8,8]
bua04	SAN52, STA22, STA02, HAL42	[-8,8]
bua05	HAL58, HAL24	[-8,8]
bua06	HAL56, HAL02	[-8,8]
bua07	HAL52, HIL22, HIL02, BAS22	[-8,8]
bua08	BAS52, BAS02	[-8,8]
bua09	CAN22, CAN02, KON22	[-8,8]
bua10	KON52, KON02	[-8,8]
bua11	UCY24, UCY22, UCY02	[-8,8]
bua12	UCY52, UCY04	[-8,8]

The first column indicates the variable name and the second column the blocks that controlled by the related variable. For instance, the variable uba01 controls UCY03 block, and take negative values if this block is occupied or allocated by a train (or trains) in Bornova direction, or positive values if the block is occupied or allocated by a train (or trains) in Ucyol direction. This variable is zero when the block is empty and is not allocated to any train. The last column is the value interval for this variable. For instance, if the number of trains that occupied or allocated the KON11 block is *n* train and their direction is Bornova uba04's value will be –*n*, from the same view point it will be +*n* if the trains' direction is Ucyol.

Train_direction_(station name)_(platform no): This variable controls platforms of station and can get negative values, positive values, and 0. The negative values between -8 and -1 indicate that the platforms are occupied or allocated by a train (or trains) in Bornova direction. The positive values between 1 and 8 point out that the platforms are occupied or allocated by a train (or trains) in Ucyol direction. The 0, means the platform is empty and not allocated to any train.

Train_arrive_(station name)_(platform no): This variable shows the arrival simulation time of a train to the platform. For Halkapinar station platform number can be 1, 2, or 3, but for other stations it can be 1 or 2.

m_(switch no): This variable controls a switch that links two parallel tracks, and can take negative values between -8 and -1, positive values between 1 and 8, and 0. The negative values indicate that the switch is occupied or allocated by a train (or trains) in Bornova direction. The positive values point outs that the switch is occupied or allocated by a train (or trains) in Ucyol direction. The 0, means the switch is empty and not allocated to any train. There are ten switches in the system. Therefore, switch no can take values from 1 to 10.

Failure_(block number): This variable denotes the situation of block. Its value may be 0 or 1. The 0 means the block is suitable for travelling. The value 1 indicates that the block is not convenient for travelling.

Train_num_getoutof: This variable demonstrates the number of trains that will leave the system after next time period begins.

Train_num_getinto: This variable denotes the number of trains that will enter the system after next time period begins.

Train_num_exist: This variable shows the number of trains that are currently in the system.

Train_num_begin: This variable demonstrates the number of trains that system needs for the first time period.

Train_num_next_period: This variable shows the number of trains that will be in the system in the next time period.

Headway_(beginning time)_(ending time): This variable indicates the headway of related time period in second. Beginning time and ending time show the time period's beginning and ending times. There are 10 time periods in weekdays, 4 time periods on Saturday, and 3 time periods on Sunday.

Carriage_number_(beginning time)_(ending time): This variable indicates the number of carriages that a train includes in related time period. Its value is always 3.

Duration_(beginning time)_(ending time): This variable indicates the duration of the related time period in second.

(station name)_getoff: This variable shows the number of passengers that alight the train at related station and leave the system.

Geton_(station name)_1: This variable denotes the number of passengers who board the train at relevant station and travel to Ucyol.

Geton_(station name)_2: This variable denotes the number of passengers who board the train at relevant station and travel to Bornova.

Waiting_(station name)_1: This variable demonstrates the number of passengers, in a certain station, waiting for the train in Ucyol direction.

Waiting_(station name)_2: This variable demonstrates the number of passengers, in a certain station, waiting for the train in Bornova direction.

RoundTripTime: It denotes the duration of a full tour, and tallies 40 minutes.

Train_waiting_time _(station name): It denotes the waiting time of the train for loading and unloading at the related station.

Preparing_time: The value of variable is 15 minutes (900 seconds), the time duration for systems preparing before the first time period.

System_open: This variable indicates that the first time period begins.

System_close: This variable indicates that the last time period ends.

Train_capacity: The value of the variable denotes the capacity of a train. We mentioned that a train consists of 3 carriages for all time periods, and a carriage has 188 passenger capacity. Thus, the capacity of a train is 564 passengers.

Train_departure_time_(end station name): This variable indicates the next departure time at an end station.

Train_ready_(end station name): The values that this variable can get are 0 and 1. The value 1 means, a train is available for departure from an end station.

Trip_number_period: This variable gives the total trips number in a time period, and calculated by a formula;

$$Trip\_number\_period = 2 * \left[\frac{Duration\_(beginningtime)\_(endingtime)}{Headway\_(beginningtime)\_(endingtime)} \right] \quad (4.2)$$

The coefficient 2 is used, because at the same time two trip begin, first is from Bornova to Ucyol, second is From Ucyol to Bornova.

Carriage_number_period: This variable denotes the total carriage number that is used for transportation for each time period, and calculated by a formula;

Carriage_number_period = 3(Trip_number_period)* (4.3)

The coefficient 3 indicates, each train consist of 3 carriages.

Tot_carriage_number: The value of this variable is calculating by summing up the *carriage_number_period* variable values for each time period, and gives the total carriage numbers that are used for transportation during a day.

Tot_passenger_number: The value of this variable demonstrates the total number of passengers who used metro for travelling during a day.

Fullness_rate: This variable indicates the fullness rate of the carriages, and calculated by a formula (3.1).

4.7 Verification and Validation of Simulation Model

After the development of the simulation model the question "does it work?" arises. The answer is found through verification and validation. Verification seeks to show that the computer program performs as expected and intended, thus providing a correct logical representation of the model. Validation on the other hand, establishes that model behaviour validly represents that of the real-word system being simulated. Both processes involve system testing that demonstrates different aspects of model accuracy (Pedgen et al., 1990, p.20).

Verification techniques such as; developing the program in a modular manner, using interactive debuggers, substituting constants for random variables and manually checking the results and animating the system are used for verifying the simulation model.

The simulation model is built in a step-by-step approach. First, the network that includes the tracks, switches and stations are modelled. Then, important functions such as; controlling the track part between departure station and destination station, beginning the trip from an end station and ending at the other end station, in addition changing the tracks are modelled. After this model is checked to see whether it works as intended, the failures and passenger arrivals to stations, train departure times from an end station, headway and number of train changes due to transition between to consecutive time periods are inserted in the model.

Interactive debugger and trace tools of Arena V2.2 software package, such as Command, Break, Trace and Watch are used for checking the status of the model, also for checking the variations in attribute and variable values.

Passenger arrivals, passenger's destination station, failure occurrence, repair times are random variables. We substitute constants for these random variables and check if passenger goes to intended destination stations, if train has limited capacity logic works, if train changes track by a switch due to failure occurrence.

Animation of the system is built by using Animate Tool of Arena V2.2 software package. In addition, the simulation model checked if it works as intended.

For validation of the model, we compare the model's performance under known conditions with the performance of the real system. As it is mentioned before, there are two responses; the average passenger time that is spent in the metro-line and the fullness rate of the carriages. There is no data about the average passenger time, but the fullness rate of the carriages is calculated as 49.5 percent.

Confidence interval is calculated for the mean value of fullness rate of carriages per week. The fullness rates of carriages per day that are obtained by simulation model, are shown in Table 4.30. The average values are indicated in last row, and are related to a week.

Table 4.30 The fullness rate of carriages per day

Week no	1	2	3	4	5	6	7	8	9	10
Monday	49.7	49.5	49.6	49.6	49.7	49.1	49.4	49.6	49.5	49.4
Tuesday	49.4	49.4	49.7	49.8	49.5	49.5	49.8	49.5	49.6	49.6
Wednesday	49.7	49.5	49.6	49.6	49.3	49.1	49.4	49.4	49.7	49.5
Thursday	49.3	49.5	49.7	49.4	49.7	49.5	49.5	49.6	49.5	49.8
Friday	49.2	49.6	49.5	49.6	50.1	49.4	50.1	49.4	49.6	49.4
Saturday	60.0	59.7	59.6	59.4	60.0	59.9	60.1	59.6	59.7	59.8
Sunday	39.3	39.3	39.4	39.1	39.2	39.1	39.3	38.8	39.6	39.5
Average	49.51	49.50	49.59	49.50	49.64	49.37	49.66	49.41	49.60	49.57

Average fullness rate values are used for calculating confidence interval. Two-side confidence interval is found by the formula;

$$\bar{x} - t_{\alpha/2,n-1} s/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1} s/\sqrt{n} \qquad (4.4)$$

Where the $\bar{x}$ and s are mean and standard deviation of sample data, n is sample size, and $t_{\alpha/2,n-1}$ is the upper $100\alpha/2$ percentage point of the t distribution with $n-1$ degrees of freedom. The related values are found as $\bar{x} = 49.54$ and $s = 0.096609$, and with 5% ($\alpha = 0.05$) significance level $t_{0,25,\ 9} = 2.262$. The lower and upper limits of confidence interval are 49.47089 and 49.60911 respectively. As it can be seen the 49.5 value is in this confidence interval (49.47089 ≤ 49.5 ≤ 49.60911) that means simulation model is valid.

CHAPTER FIVE
RSM STUDY

The simulation model of the system is built by using Arena V2.2 software package. It is a generic model that easily changes to suit the changes in headways, in the length of time periods, and in the number of carriages.

MINITAB Release 13.20 software package is used for building two-level full factorial designs, constructing normal probability plot of the residuals, residuals versus predicted response plot and autocorrelation diagram, analysing variance, obtaining first order regression models, building *CCD*s, obtaining second order regression models, and running Derringer-Suich multi-response optimization procedure.

Because of software limitation (MINITAB can only build a *CCD* and analyze a second order model with maximum 6 independent factors), the model which is for weekdays from Monday to Friday is divided into two parts. Therefore, there are four models, the first one is for weekdays morning with five time periods, second is for weekdays afternoon with five time periods, the next is for Saturday with four time periods, and the last one is for Sunday with three time periods. The beginning and ending times of each time period are denoted in Table 5.1.

Table 5.1 Related time periods for four separate models

Monday-Friday		Saturday	Sunday
Morning	**Afternoon**		
06:00 – 07:00	11:30 - 17:00	06:00 - 11:00	06:00 - 09:00
07:00 – 07:30	17:00 - 18:30	11:00 - 19:00	09:00 - 20:00
07:30 – 09:00	18:30 - 19:00	19:00 - 22:00	20:00 - 00:00
09:00 – 09:30	19:00 - 22:00	22:00 - 00:00	
09:30 – 11:30	22:00 - 00:00		

As mentioned before, the aim of this study is to find the headways (input factors) minimizing the average passenger time spent in the metro-line (first response) with the requirement as the fullness rate (second response), fifty percent, of the carriages.

5.1 Weekday Morning Problem (*WMP*)

5.1.1 Estimation Process

5.1.1.1 Phase Zero

The objective of *WMP* is to find the levels of 5 input factors that related to the 5 time periods from 06:00 a.m. to 11:30 a.m. and which minimize the average passenger time spent in the metro-line with the requirement as the fifty percent fullness rate of the carriages. The built simulation model is explained in chapter four.

Generally at phase zero, a screening experiment is made for investigating potential input factors, which are thought to be important in the response surface study, and for determining important factors. Because the input factors are 5 headways and none of them can be eliminated we skip factor screening processes.

The input factors are respectively;

X_1: The headway for the first time period from 06:00 a.m. to 07:00 a.m.

X_2: The headway for the second time period from 07:00 a.m. to 07:30 a.m.

X_3: The headway for the third time period from 07:30 a.m. to 09:00 a.m.

X_4: The headway for the fourth time period from 09:00 a.m. to 09:30 a.m.

X_5: The headway for the fifth time period from 09:30 a.m. to 11:30 a.m.

The responses are respectively;

Y_1: The average passenger time that is spent in the metro-line (in second)

Y_2: The fullness rate of the carriages (in percentage)

5.1.1.2 Phase One

Low and high level of input factors for *WMP*, time period durations, and the natural values are given in Table 5.2. Coded values are found by using formula (2.4), as an example the low-level of X_1 is calculated as;

$$X_1 = \frac{300-[600+300]/2}{[600-300]/2} = -1$$

Table 5.2 Low and high level of input factors for *WMP*

Input factor	Time period	Duration (hour)	(second)	Low level Coded	Natural (second)	High level Coded	Natural (second)
X_1	06:00 - 07:00	1,0	3600	-1	300	1	600
X_2	07:00 - 07:30	0,5	1800	-1	300	1	600
X_3	07:30 - 09:00	1,5	5400	-1	300	1	600
X_4	09:00 - 09:30	0,5	1800	-1	300	1	600
X_5	09:30 - 11:30	2,0	7200	-1	300	1	600

The first order regression models (*form*) with two-factor interactions are assumed to be;

$$Y_1 = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_{12}X_1X_2 + \beta_{13}X_1X_3 + \beta_{14}X_1X_4 + \beta_{15}X_1X_5 + \beta_{23}X_2X_3 + \beta_{24}X_2X_4 + \beta_{25}X_2X_5 + \beta_{34}X_3X_4 + \beta_{35}X_3X_5 + \beta_{45}X_4X_5 + \varepsilon$$

$$Y_2 = \delta_0 + \delta_1X_1 + \delta_2X_2 + \delta_3X_3 + \delta_4X_4 + \delta_5X_5 + \delta_{12}X_1X_2 + \delta_{13}X_1X_3 + \delta_{14}X_1X_4 + \delta_{15}X_1X_5 + \delta_{23}X_2X_3 + \delta_{24}X_2X_4 + \delta_{25}X_2X_5 + \delta_{34}X_3X_4 + \delta_{35}X_3X_5 + \delta_{45}X_4X_5 + \varepsilon \quad (5.1)$$

where β_0 and δ_0 are constant, and β_1, β_2, β_3, β_4, β_5, δ_1, δ_2, δ_3, δ_4 and δ_5 are coefficients corresponding to main effects, and β_{12}, β_{13}, β_{14}, β_{15}, β_{23}, β_{24}, β_{25}, β_{34}, β_{35}, β_{45}, δ_{12}, δ_{13}, δ_{14}, δ_{15}, δ_{23}, δ_{24}, δ_{25}, δ_{34}, δ_{35} and δ_{45} are coefficients corresponding to two-factor interaction effects and ε is statistical error that have a normal distribution with mean zero and variance σ^2.

5.1.1.2.1 Two-level Full Factorial Design

For fitting a first order regression model a two-level full factorial design (2^5) with central runs is designed and then the simulation model is run 10 times at each design point, and response values for each design points are found.

Although the values for two-level full factorial design points, which are denoted in Table 5.3, are average of 10 replications, the values related to the central points are obtained by one replication.

Table 5.3 Simulation results for *WMP* (2^5 design with 5 central runs)

	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2
1	-1	-1	-1	-1	-1	525	31,1
2	1	-1	-1	-1	-1	531	34,0
3	-1	1	-1	-1	-1	537	32,4
4	1	1	-1	-1	-1	542	36,0
5	-1	-1	1	-1	-1	595	35,9
6	1	-1	1	-1	-1	602	40,1
7	-1	1	1	-1	-1	603	37,7
8	1	1	1	-1	-1	605	42,4
9	-1	-1	-1	1	-1	555	32,3
10	1	-1	-1	1	-1	561	35,8
11	-1	1	-1	1	-1	567	34,1
12	1	1	-1	1	-1	571	37,8
13	-1	-1	1	1	-1	611	37,8
14	1	-1	1	1	-1	617	42,6
15	-1	1	1	1	-1	615	40,0
16	1	1	1	1	-1	619	45,3
17	-1	-1	-1	-1	1	592	37,9
18	1	-1	-1	-1	1	597	42,4
19	-1	1	-1	-1	1	604	40,0
20	1	1	-1	-1	1	608	45,3
21	-1	-1	1	-1	1	663	45,1
22	1	-1	1	-1	1	668	52,0
23	-1	1	1	-1	1	670	48,3
24	1	1	1	-1	1	672	56,4
25	-1	-1	-1	1	1	611	39,9
26	1	-1	-1	1	1	617	45,0
27	-1	1	-1	1	1	623	42,5
28	1	1	-1	1	1	627	48,7
29	-1	-1	1	1	1	666	48,3
30	1	-1	1	1	1	671	56,2
31	-1	1	1	1	1	672	51,9
32	1	1	1	1	1	674	61,2
33	0	0	0	0	0	611	46,2
34	0	0	0	0	0	611	46,1
35	0	0	0	0	0	611	46,4
36	0	0	0	0	0	610	46,0
37	0	0	0	0	0	611	46,4

For verifying that the least squares regression assumptions are not violated, the residuals from the least squares fit are checked to see whether they are normally, identically and independently distributed with zero mean and constant variance.

The residuals are shown in Table 5.4, are calculated by the formula;

$$e_i = Y_i - \hat{Y}_i \qquad (5.2)$$

where $\hat{Y}_i$ is predicted response.

Table 5.4 Residuals for Y_1 response for *WMP*

1	2	3	4	5	6	7	8	9	10
0,27	-0,29	-0,54	0,27	-1,29	0,02	0,52	-0,79	-0,29	-0,98
11	12	13	14	15	16	17	18	19	20
0,02	-0,29	0,27	0,46	-0,79	-0,23	0,21	-0,48	-0,98	-0,29
21	22	23	24	25	26	27	28	29	30
-0,23	-0,04	0,21	-0,23	-0,23	-0,04	-0,29	0,27	-0,54	-0,48
31	32	33	34	35	36	37			
0,02	-0,54	1,68	1,68	1,68	0,68	1,68			

In this study, we use a normal probability plot of the residuals to check normality assumption. The normal probability plot of the residuals for Y_1 response is shown in Figure 5.1. As can be seen, the residuals plot is approximately along a straight line. Thus we conclude that the normality assumption is satisfied.

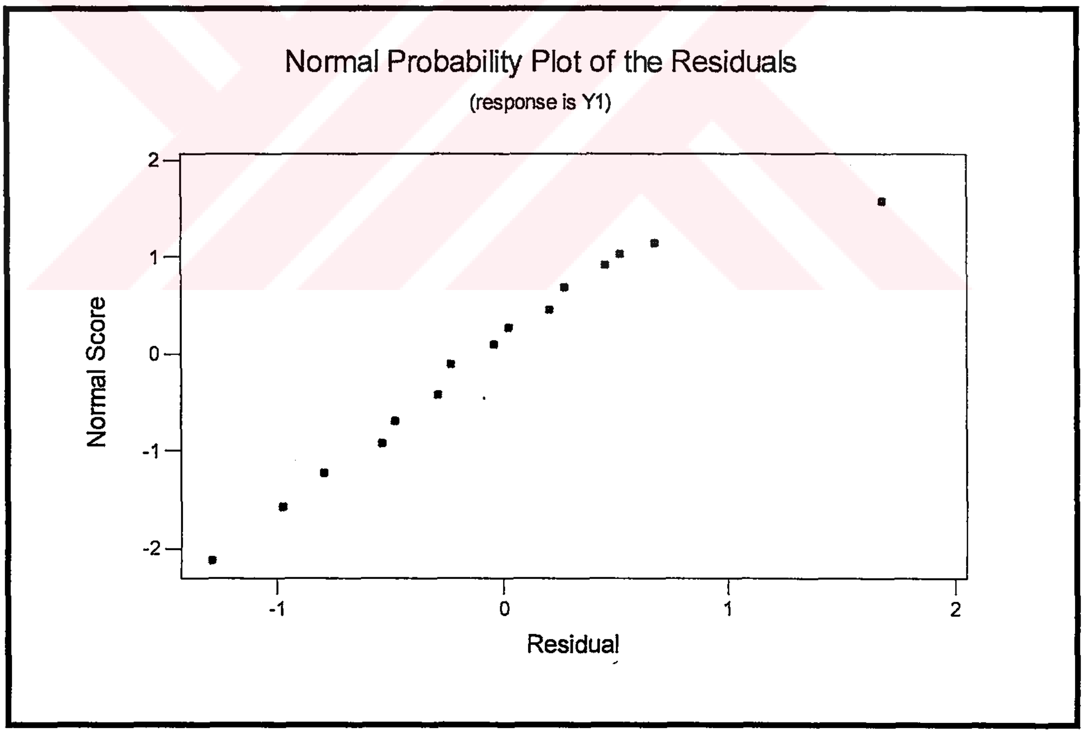

Figure 5.1 Normal probability plot of the residuals for Y_1 response for *WMP*

The plot of the residuals versus predicted response is also demonstrated in Figure 5.2. As can be seen, the residuals scatter randomly on the plot that indicates the variance is constant for all values of the response.

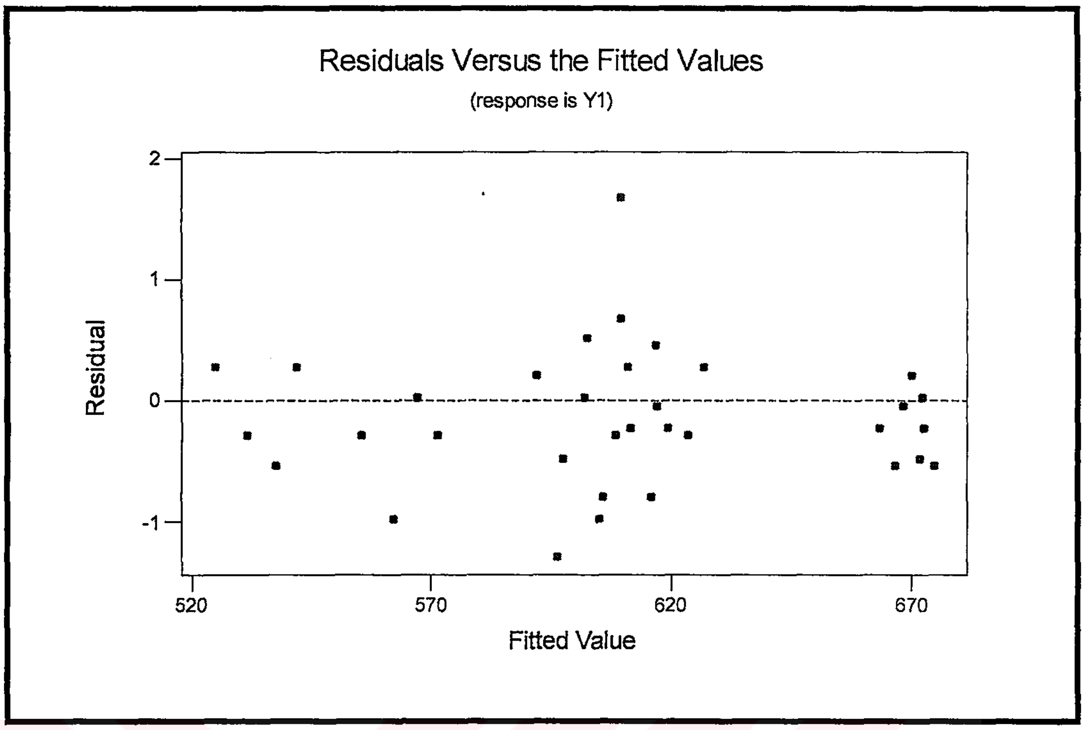

Figure 5.2 Residuals versus predicted response plot for Y_1 response for *WMP*

The assumption that the residuals are independent is controlled by autocorrelation diagram. As can be seen from the Figure 5.3, the assumption is not violated.

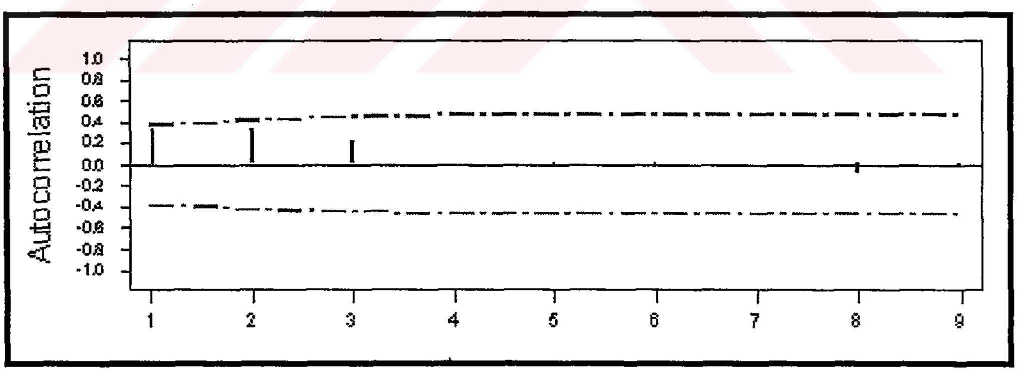

Figure 5.3 Autocorrelation diagram for Y_1 response for *WMP*

For the zero mean assumption a hypothesis test on the mean of the residuals with unknown variance is built. The null and alternative hypotheses are respectively;

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0 \quad (5.3)$$

To test the $H_0: \mu = \mu_0$, the value of the test statistic t_0 is calculated by the formula;

$$t_0 = \frac{\overline{X} - \mu_0}{s / \sqrt{n}} \tag{5.4}$$

where $\overline{X}$ is the sample mean, s is the sample standard deviation, and n is the sample size. H_0 is rejected if $t_0 > t_{\alpha/2,n-1}$ or $t_0 < -t_{\alpha/2,n-1}$, where $t_{\alpha/2,n-1}$ and $-t_{\alpha/2,n-1}$ are the upper and lower 100α/2 percentage points of the t with $n-1$ degrees of freedom.

The values for the calculation are $\overline{X} = 0.00032$, $\mu_0 = 0$, $s = 0.73389$, and $n = 37$, and with 5% ($\alpha = 0.05$) significance level $t_{0,25,\ 36} = 2{,}0294$, $-t_{0.25,\ 36} = -2.0294$. Substituting the values in formula (5.4) t_0 is found as 0.002652.

Since t_0 is between -2.0294 and 2.0294 we can not reject that the mean of the residuals is zero hypothesis.

After verifying that the residuals are normally, identically and independently distributed with zero mean and constant variance for the Y_2 response, Variance Analysis is used to see if the main effects, two-factor interactions, curvature and lack-of-fit are statistically significant. Table 5.5 is the Analysis of Variance Table (ANOVA) for the first response Y_1, and Table 5.6 for the second response Y_2. In addition, estimated effects and coefficients (*EEC*) for two responses are given in Table 5.7 and 5.8, respectively.

In Table 5.5, *DF* denotes degrees of freedom, *Seq SS*; sequential sum of squares, *Adj SS*; adjusted sum of squares, *Adj MS*; adjusted mean square, *F*;is the statistic F, and *P*; the P-value, *S*; standard deviation, *R-sq*; and *R-sq(adj)*; adjusted coefficient of determination.

Table 5.5 ANOVA table for Y_1 response for *form* for *WMP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	61284,2	61284,2	12256,8	10000,00	0,000
2-Way Interactions	10	872,6	872,6	87,3	94,50	0,000
Residual Error	21	19,4	19,4	0,9		
Curvature	1	12,6	12,6	12,6	37,03	0,000
Lack of Fit	16	6,0	6,0	0,4	1,87	0,287
Pure Error	4	0,8	0,8	0,2		
Total	36	62176,1				
S = 5,364 R-Sq = 98,6% R-Sq(adj) = 98,3%						

The null and alternative hypotheses related to main effects (5.5), two-way interactions (5.6), curvature (5.7) and lack-of-fit (5.8) for Variance Analysis are;

$H_0 : \beta_1=\beta_2=\beta_3=\beta_4=\beta_5=0$

$H_1 : \beta_i \neq 0$ *for at least one i (i=1,2,3,4,5)* (5.5)

$H_0 : \beta_{12}=\beta_{13}=\beta_{14}=\beta_{15}=\beta_{23}=\beta_{24}=\beta_{25}=\beta_{34}=\beta_{35}=\beta_{45}=0$

$H_1 : \beta_{ij} \neq 0$ *for at least one ij (i=1,2,3,4 and j=2,3,4,5)* (5.6)

$H_0 : \beta_{11}=\beta_{22}=\beta_{33}=\beta_{44}=\beta_{55}=0$

$H_1 : \beta_{ii} \neq 0$ *for at least one i (i=1,2,3,4,5)* (5.7)

H_0 *: The regression model is correct*

H_1 *: The regression model is not correct* (5.8)

where β_1, β_2, β_3, β_4 and β_5 are coefficients corresponding to main effects, and β_{12}, β_{13}, β_{14}, β_{15}, β_{23}, β_{24}, β_{25}, β_{34}, β_{35} and β_{45} are coefficients corresponding to two-factor interaction effects, β_{11}, β_{22}, β_{33}, β_{44} and β_{55} are coefficients corresponding to square effects.

In Table 5.5, since the *P*-value for the lack-of-fit test is greater than the significance level ($\alpha = 0.05$), the lack-of-fit is not statistically significant [H_0 (5.8) is not rejected], that is, the first order regression model for Y_1 response for *WMP* is adequate. The main effects [H_0 (5.5) is rejected], two-way interactions [H_0 (5.6) is rejected] and curvature [H_0 (5.7) is rejected] are statistically significant (*P*-value<0.05).

Table 5.6 ANOVA table for Y_2 response for *form* for *WMP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	1701,9	1701,9	340,4	108,95	0,000
2-Way Interactions	10	62,3	62,3	6,2	1,99	0,088
Residual Error	21	65,6	65,6	3,1		
Curvature	1	63,5	63,5	63,5	607,30	0,000
Lack of Fit	16	2,0	2,0	0,1	3,84	0,101
Pure Error	4	0,1	0,1	0,0		
Total	36	1829,8				
S = 2,031 R-Sq = 93,0% R-Sq(adj) = 91,9%						

Table 5.6 shows that the lack-of-fit and the two-way interactions are not statistically significant (*P*>0.05), curvature and main effects are statistically significant (*P*<0.05).

In Table 5.7, *SE Coef* is standard error of coefficient, *T* is the statistic T, and *P* is the P-value.

Table 5.7 *EEC* table for Y_1 response for *form* for *WMP*

Term	Effect	Coef	SE Coef	T	P
Constant		609,324	0,1580	3857,25	0,000
X1	4,562	2,281	0,1699	13,43	0,000
X2	7,937	3,969	0,1699	23,36	0,000
X3	59,687	29,844	0,1699	175,69	0,000
X4	16,437	8,219	0,1699	48,38	0,000
X5	61,188	30,594	0,1699	180,11	0,000
*X1*X2*	-1,188	-0,594	0,1699	-3,50	0,002
*X1*X3*	-0,438	-0,219	0,1699	-1,29	0,212
*X1*X4*	0,063	0,031	0,1699	0,18	0,856
*X1*X5*	-0,438	-0,219	0,1699	-1,29	0,212
*X2*X3*	-3,313	-1,656	0,1699	-9,75	0,000
*X2*X4*	-0,562	-0,281	0,1699	-1,66	0,113
*X2*X5*	0,188	0,094	0,1699	0,55	0,587
*X3*X4*	-8,063	-4,031	0,1699	-23,73	0,000
*X3*X5*	-0,063	-0,031	0,1699	-0,18	0,856
*X4*X5*	-5,563	-2,781	0,1699	-16,37	0,000

Table 5.8 ***EEC*** **table for** Y_2 **response for** ***form*** **for** ***WMP***

Term	Effect	Coef	SE Coef	T	P
Constant		42,905	0,2906	147,65	0,000
X1	5,375	2,688	0,3125	8,60	0,000
X2	2,725	1,363	0,3125	4,36	0,000
X3	7,875	3,938	0,3125	12,60	0,000
X4	2,650	1,325	0,3125	4,24	0,000
X5	10,363	5,181	0,3125	16,58	0,000
*X1*X2*	0,400	0,200	0,3125	0,64	0,529
*X1*X3*	1,025	0,513	0,3125	1,64	0,116
*X1*X4*	0,350	0,175	0,3125	0,56	0,581
*X1*X5*	1,288	0,644	0,3125	2,06	0,052
*X2*X3*	0,425	0,213	0,3125	0,68	0,504
*X2*X4*	0,225	0,113	0,3125	0,36	0,722
*X2*X5*	0,713	0,356	0,3125	1,14	0,267
*X3*X4*	0,525	0,263	0,3125	0,84	0,410
*X3*X5*	1,838	0,919	0,3125	2,94	0,008
*X4*X5*	0,638	0,319	0,3125	1,02	0,319

As a result, the fitted first order models for both responses are as follows;

$$Y_1 = 609{,}324 + 2{,}281\,x_1 + 3{,}969\,x_2 + 29{,}844\,x_3 + 8{,}219\,x_4 + 30{,}594\,x_5 - 0{,}594\,x_1x_2 - 0{,}219\,x_1x_3 + 0{,}031\,x_1x_4 - 0{,}219\,x_1x_5 - 1{,}656\,x_2x_3 - 0{,}281\,x_2x_4 + 0{,}094\,x_2x_5 - 4{,}031\,x_3x_4 - 0{,}031\,x_3x_5 - 2{,}781\,x_4x_5$$

$$Y_2 = 42{,}905 + 2{,}688\,x_1 + 1{,}363\,x_2 + 3{,}938\,x_3 + 1{,}325\,x_4 + 5{,}181\,x_5 + 0{,}200\,x_1x_2 + 0{,}513\,x_1x_3 + 0{,}175\,x_1x_4 + 0{,}644\,x_1x_5 + 0{,}213\,x_2x_3 + 0{,}113\,x_2x_4 + 0{,}356\,x_2x_5 + 0{,}263\,x_3x_4 + 0{,}919\,x_3x_5 + 0{,}319\,x_4x_5 \quad (5.9)$$

Because the true response surface usually exhibits curvature near the optimum, it is understood that the determined region of experimentation is near the region of optimum. As mentioned in chapter two, the second phase of *RSM* begins. Therefore, second order models can be built.

5.1.2 Optimization Process

5.1.2.1 Phase Two

5.1.2.1.1 *CCD* and Development of Metamodels for *WMP*

The second-order regression models (*sorm*) are assumed to be;

$$Y_1 = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_{11}X_1^2 + \beta_{22}X_2^2 + \beta_{33}X_3^2 + \beta_{44}X_4^2 + \beta_{55}X_5^2 + \beta_{12}X_1X_2 + \beta_{13}X_1X_3 + \beta_{14}X_1X_4 + \beta_{15}X_1X_5 + \beta_{23}X_2X_3 + \beta_{24}X_2X_4 + \beta_{25}X_2X_5 + \beta_{34}X_3X_4 + \beta_{35}X_3X_5 + \beta_{45}X_4X_5 + \varepsilon$$

$$Y_2 = \delta_0 + \delta_1X_1 + \delta_2X_2 + \delta_3X_3 + \delta_4X_4 + \delta_5X_5 + \delta_{11}X_1^2 + \delta_{22}X_2^2 + \delta_{33}X_3^2 + \delta_{44}X_4^2 + \delta_{55}X_5^2 + \delta_{12}X_1X_2 + \delta_{13}X_1X_3 + \delta_{14}X_1X_4 + \delta_{15}X_1X_5 + \delta_{23}X_2X_3 + \delta_{24}X_2X_4 + \delta_{25}X_2X_5 + \delta_{34}X_3X_4 + \delta_{35}X_3X_5 + \delta_{45}X_4X_5 + \varepsilon \quad (5.10)$$

where β_0 and δ_0 are constant, and β_1, β_2, β_3, β_4, β_5, δ_1, δ_2, δ_3, δ_4 and δ_5 are coefficients corresponding to main effects, and β_{12}, β_{13}, β_{14}, β_{15}, β_{23}, β_{24}, β_{25}, β_{34}, β_{35}, β_{45}, δ_{12}, δ_{13}, δ_{14}, δ_{15}, δ_{23}, δ_{24}, δ_{25}, δ_{34}, δ_{35} and δ_{45} are coefficients corresponding to two-factor interaction effects, β_{11}, β_{22}, β_{33}, β_{44}, β_{55}, δ_{11}, δ_{22}, δ_{33}, δ_{44} and δ_{55} are coefficients corresponding to square effects, and ε is statistical error that have a normal distribution with mean zero and variance σ^2.

For fitting second order models a *CCD* is built by augmenting the two-level full factorial design with central runs and axial runs ($\alpha = 1$, means design is face-centered). In Table 5.9, the *CCD* used for fitting second order models for two responses is given. The values for two-level full factorial design points and axial points are the average of 10 replications, and the values related to the central points are obtained by one replication.

After verifying that the residuals of the second order regression models fitted for Y_1 and Y_2 are normally, identically and independently distributed with zero mean and constant variance, the ANOVA tables are given in Table 5.10 and Table 5.11. Also estimated effects and coefficients (*EEC*) for two responses are denoted in Table 5.12 and 5.13 respectively.

Table 5.9 Simulation results for *WMP* (*CCD*)

	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2
1	-1	-1	-1	-1	-1	525	31,1
2	1	-1	-1	-1	-1	531	34,0
3	-1	1	-1	-1	-1	537	32,4
4	1	1	-1	-1	-1	542	36,0
5	-1	-1	1	-1	-1	595	35,9
6	1	-1	1	-1	-1	602	40,1
7	-1	1	1	-1	-1	603	37,7
8	1	1	1	-1	-1	605	42,4
9	-1	-1	-1	1	-1	555	32,3
10	1	-1	-1	1	-1	561	35,8
11	-1	1	-1	1	-1	567	34,1
12	1	1	-1	1	-1	571	37,8
13	-1	-1	1	1	-1	611	37,8
14	1	-1	1	1	-1	617	42,6
15	-1	1	1	1	-1	615	40,0
16	1	1	1	1	-1	619	45,3
17	-1	-1	-1	-1	1	592	37,9
18	1	-1	-1	-1	1	597	42,4
19	-1	1	-1	-1	1	604	40,0
20	1	1	-1	-1	1	608	45,3
21	-1	-1	1	-1	1	663	45,1
22	1	-1	1	-1	1	668	52,0
23	-1	1	1	-1	1	670	48,3
24	1	1	1	-1	1	672	56,4
25	-1	-1	-1	1	1	611	39,9
26	1	-1	-1	1	1	617	45,0
27	-1	1	-1	1	1	623	42,5
28	1	1	-1	1	1	627	48,7
29	-1	-1	1	1	1	666	48,3
30	1	-1	1	1	1	671	56,2
31	-1	1	1	1	1	672	51,9
32	1	1	1	1	1	674	61,2
33	-1	0	0	0	0	607	42,4
34	1	0	0	0	0	612	48,5
35	0	-1	0	0	0	606	44,1
36	0	1	0	0	0	615	47,4
37	0	0	-1	0	0	579	40,8
38	0	0	1	0	0	641	49,4
39	0	0	0	-1	0	607	44,3
40	0	0	0	1	0	625	47,3
41	0	0	0	0	-1	579	39,3
42	0	0	0	0	1	640	50,8
43	0	0	0	0	0	610	46,5
44	0	0	0	0	0	607	46,1
45	0	0	0	0	0	611	46,8
46	0	0	0	0	0	609	46,5
47	0	0	0	0	0	610	45,5
48	0	0	0	0	0	608	45,8
49	0	0	0	0	0	610	45,7
50	0	0	0	0	0	612	46,2
51	0	0	0	0	0	614	45,9
52	0	0	0	0	0	610	46,2

Table 5.10 ANOVA table for Y_1 response for *sorm* for *WMP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	20	66235,5	66235,5	3311,8	2000,00	0,000
Linear	5	65277,4	65277,4	13055,5	8000,00	0,000
Square	5	85,6	85,6	17,1	9,84	0,000
Interaction	10	872,6	872,6	87,3	50,15	0,000
Residual Error	31	53,9	53,9	1,7		
Lack-of-Fit	22	19,0	19,0	0,9	0,22	0,998
Pure Error	9	34,9	34,9	3,9		
Total	51	66289,4				
S = 1,319 R-Sq = 99,9% R-Sq(adj) = 99,9%						

Table 5.11 ANOVA table for Y_2 response for *sorm* for *WMP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	20	2039,1	2039,1	102,0	662,00	0,000
Linear	5	1832,2	1832,2	366,4	2000,00	0,000
Square	5	144,6	144,6	28,9	187,69	0,000
Interaction	10	62,3	62,3	6,2	40,44	0,000
Residual Error	31	4,8	4,8	0,2		
Lack-of-Fit	22	3,3	3,3	0,2	0,91	0,594
Pure Error	9	1,5	1,5	0,2		
Total	51	2043,9				
S = 0,3925 R-Sq = 99,8% R-Sq(adj) = 99,6%						

In Table 5.10 and Table 5.11, we see that linear, quadratic and two-way interaction effects are statistically significant ($P<0.05$), and the second order regression models are statistically significant ($P>0.05$) for both responses.

As a result, the second order regression models (metamodels) for Y_1 and Y_2 are respectively;

$$Y_1 = 610,682 + 2,294\,x_1 + 4,000\,x_2 + 29,912\,x_3 + 8,265\,x_4 + 30,588\,x_5 - 1,908\,x_1^2 - 0,908\,x_2^2 - 1,408\,x_3^2 + 4,592\,x_4^2 - 1,908\,x_5^2 - 0,594\,x_1x_2 - 0,219\,x_1x_3 + 0,031\,x_1x_4 - 0,219\,x_1x_5 - 1,656\,x_2x_3 - 0,281\,x_2x_4 + 0,094\,x_2x_5 - 4,031\,x_3x_4 - 0,031\,x_3x_5 - 2,781\,x_4x_5$$

$$Y_2 = 46,147 + 2,709\,x_1 + 1,379\,x_2 + 3,959\,x_3 + 1,335\,x_4 + 5,215\,x_5 - 0,732\,x_1^2 - 0,432\,x_2^2 - 1,082\,x_3^2 - 0,382\,x_4^2 - 1,132\,x_5^2 + 0,200\,x_1x_2 + 0,513\,x_1x_3 + 0,175\,x_1x_4 + 0,644\,x_1x_5 + 0,213\,x_2x_3 + 0,113\,x_2x_4 + 0,356\,x_2x_5 + 0,263\,x_3x_4 + 0,919\,x_3x_5 + 0,319\,x_4x_5 \quad (5.11)$$

Table 5.12 *EEC* table for Y_1 response for *sorm* for *WMP*

Term	Coef	SE Coef	T	P
Constant	610,682	0,327	1870,3860	0,000
X1	2,294	0,226	10,1410	0,000
X2	4,000	0,226	17,6820	0,000
X3	29,912	0,226	132,2280	0,000
X4	8,265	0,226	36,5350	0,000
X5	30,588	0,226	135,2190	0,000
*X1*X1*	-1,908	0,838	-2,2770	0,030
*X2*X2*	-0,908	0,838	-1,0840	0,287
*X3*X3*	-1,408	0,838	-1,6800	0,103
*X4*X4*	4,592	0,838	5,4770	0,000
*X5*X5*	-1,908	0,838	-2,2770	0,030
*X1*X2*	-0,594	0,233	-2,5460	0,016
*X1*X3*	-0,219	0,233	-0,9380	0,355
*X1*X4*	0,031	0,233	0,1340	0,894
*X1*X5*	-0,219	0,233	-0,9380	0,355
*X2*X3*	-1,656	0,233	-7,1030	0,000
*X2*X4*	-0,281	0,233	-1,2060	0,237
*X2*X5*	0,094	0,233	0,4020	0,690
*X3*X4*	-4,031	0,233	-17,2890	0,000
*X3*X5*	-0,031	0,233	-0,1340	0,894
*X4*X5*	-2,781	0,233	-11,9280	0,000

Table 5.13 *EEC* table for Y_2 response for *sorm* for *WMP*

Term	Coef	SE Coef	T	P
Constant	46,147	0,097	474,9770	0,000
X1	2,709	0,067	40,2410	0,000
X2	1,379	0,067	20,4920	0,000
X3	3,959	0,067	58,8110	0,000
X4	1,335	0,067	19,8370	0,000
X5	5,215	0,067	77,4680	0,000
*X1*X1*	-0,732	0,249	-2,9320	0,006
*X2*X2*	-0,432	0,249	-1,7300	0,094
*X3*X3*	-1,082	0,249	-4,3360	0,000
*X4*X4*	-0,382	0,249	-1,5290	0,136
*X5*X5*	-1,132	0,249	-4,5360	0,000
*X1*X2*	0,200	0,069	2,8820	0,007
*X1*X3*	0,513	0,069	7,3860	0,000
*X1*X4*	0,175	0,069	2,5220	0,017
*X1*X5*	0,644	0,069	9,2780	0,000
*X2*X3*	0,213	0,069	3,0630	0,005
*X2*X4*	0,113	0,069	1,6210	0,115
*X2*X5*	0,356	0,069	5,1340	0,000
*X3*X4*	0,263	0,069	3,7830	0,001
*X3*X5*	0,919	0,069	13,2410	0,000
*X4*X5*	0,319	0,069	4,5940	0,000

After building regression metamodels for both responses, they are verified and validated.

5.1.2.1.2 Verification and Validation of *WMP* Metamodels

For verification, lack-of-fit tests are applied for both metamodels. As can be seen from Table 5.10 and Table 5.11, the metamodels have no statistically significant lack-of-fit (P>0.05) for both responses with 5% significance level ($\alpha = 0.05$). R^2 is 99,9%, and R^2*-adjusted* is 99.9% for response Y_1, and R^2 is 99.8%, and R^2*-adjusted* is 99.6% for response Y_2.

For validation, 32 (2^5) points (because there are 5 input factors with 2 level and we want to see if metamodel is valid in the entire experimentation region) that selected randomly in experimentation region, and from different design points, are used. These 32 points are shown in Table 5.14. The simulation model is run at these randomly selected points with 10 replications for each point. The values that are denoted in the Simulation Model column of Table 5.14 are the average of these 10 replications. Also the values denoted in Metamodel column are predicted by fitted second order regression metamodels.

The Absolute Relative Error (*ARE*) is selected as a criterion, *ARE* value (r) is calculated by the following formula;

$$r = \left|(w - y) / w\right| \qquad (5.12)$$

where simulation output is denoted by w, and metamodel output is denoted by y. In addition, a threshold (r_{max}) is quantified as 3% (Kleijnen & Sargent, 2000, p.20).

Since the entire *ARE* values are smaller than 3% the second order metamodels can be used for prediction.

Table 5.14 Metamodels validation for *WMP*

	Input factors					Simulation model		Metamodel		ARE (%)	
	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
1	-0,6	-0,2	-0,6	-0,2	-0,6	564,0	38,10	567,6	38,32	0,63	0,57
2	0,2	-0,2	-0,2	-0,6	-0,6	574,0	41,00	580,3	41,25	1,10	0,61
3	-0,6	0,2	-0,6	-0,2	-0,6	568,0	38,30	569,7	38,68	0,30	0,98
4	0,6	0,6	-0,2	-0,6	-0,2	592,0	45,30	597,2	45,19	0,88	0,25
5	-0,2	-0,2	0,6	-0,6	-0,6	609,0	42,50	605,0	42,56	0,65	0,14
6	0,2	-0,6	0,6	-0,2	-0,6	602,0	43,60	606,1	43,54	0,68	0,14
7	-0,2	0,6	0,2	-0,6	-0,2	626,0	44,90	608,8	44,88	2,75	0,03
8	0,2	0,2	0,2	-0,2	-0,6	596,0	43,60	597,0	43,67	0,17	0,16
9	-0,6	-0,6	-0,2	0,6	-0,6	597,0	39,70	588,5	39,95	1,42	0,62
10	0,6	-0,6	-0,6	0,2	-0,2	583,0	42,80	586,2	42,78	0,55	0,04
11	-0,2	0,2	-0,6	0,6	-0,2	598,0	42,60	594,8	42,66	0,53	0,13
12	0,2	0,2	-0,2	0,6	-0,6	590,0	43,00	594,9	43,07	0,83	0,16
13	-0,2	-0,2	0,2	0,6	-0,6	605,0	42,90	603,4	43,03	0,27	0,30
14	0,2	-0,2	0,6	0,2	-0,6	607,0	44,80	610,6	44,73	0,59	0,16
15	-0,6	0,6	0,6	0,2	-0,2	619,0	45,90	623,0	45,84	0,65	0,14
16	0,6	0,6	0,6	0,2	-0,6	610,0	46,60	612,7	46,55	0,44	0,10
17	-0,6	-0,6	-0,2	-0,6	0,2	604,0	42,60	601,9	42,60	0,34	0,00
18	0,6	-0,2	-0,6	-0,6	0,6	600,0	45,60	605,8	45,84	0,97	0,53
19	-0,2	0,6	-0,6	-0,2	0,2	606,0	44,00	598,7	44,08	1,21	0,18
20	0,2	0,2	-0,2	-0,6	0,6	615,0	47,80	620,7	47,74	0,93	0,13
21	-0,6	-0,2	0,6	-0,2	0,2	633,0	46,50	630,6	46,48	0,38	0,04
22	0,6	-0,6	0,2	-0,2	0,6	634,0	50,10	631,5	49,84	0,40	0,53
23	-0,2	0,6	0,6	-0,6	0,2	631,0	48,50	633,8	48,26	0,44	0,50
24	0,6	0,6	0,2	-0,2	0,2	629,0	50,10	623,8	49,90	0,83	0,39
25	-0,6	-0,2	-0,6	0,6	0,6	605,0	44,10	613,8	44,16	1,46	0,13
26	0,6	-0,2	-0,2	0,2	0,2	612,0	47,70	612,4	47,60	0,07	0,21
27	-0,2	0,2	-0,6	0,6	0,6	619,0	46,20	617,4	46,13	0,26	0,15
28	0,2	0,2	-0,2	0,6	0,2	617,0	48,20	618,6	47,77	0,26	0,90
29	-0,2	-0,6	0,6	0,2	0,2	632,0	48,00	632,7	47,78	0,11	0,46
30	0,2	-0,6	0,6	0,2	0,6	639,0	51,40	645,1	50,94	0,95	0,89
31	-0,6	0,6	0,2	0,2	0,6	635,0	48,70	635,7	48,59	0,12	0,22
32	0,6	0,2	0,2	0,2	0,6	635,0	52,20	636,8	52,04	0,29	0,30

After verification and validation of regression metamodels, Derringer-Suich multi-response optimization procedure is used for optimization.

5.1.2.1.3 Derringer-Suich Multi-Response Optimization Procedure for *WMP*

As mentioned in chapter two, Derringer-Suich method uses a desirability function in which the researcher's priorities and desires on the response values are built into one optimization procedure. Firstly, researcher must make a decision about upper limit, lower limit and target values of each response. In the current problem, since the target value for the fullness rate of the carriages response (Y_2) is 50%, the lower and upper limits are determined as 45% and 55%, respectively. To decide the upper and lower limits on the average passenger time spent in the metro-line response (Y_1) the simulation model is run 20 times at the low levels of all input factors. The lower limit is defined as the average of these 20 values, and is found as 525 seconds. To determine the upper limit, the simulation model is run 20 times at the high levels of all input factors, and the average value, 675 seconds, is found as upper limit.

Also the t, and s weight values must be determined. We will use three values 0.1, 1 and 10 for t and s. Therefore, nine situations (combination of weights) appear to be evaluated.

The individual desirability values and the composite desirability values for each design point are given in Table 5.15. The individual desirability values are found by using the formulas (2.34) and (2.36), and the composite desirability values are found by using the formula (2.37). The response values are predicted by using related second order regression metamodels. Response Optimiser tool of MINITAB software package is used for finding global optimum points for each combination of weights.

For each combination of weights, the lower, target, upper values and weights for two responses, and the optimum conditions with coded variables, and also corresponding predicted responses at these coded values which are calculated from the fitted models, in addition the individual desirability values and the composite desirability values at optimum conditions are shown in Table 5.16.

Table 5.15 Individual and composite desirability values for *WMP* ($s = t = 0.1$)

	Input factors					Simulation model		Metamodel		Individual desirability		Composite desirability
	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2	Y_1	Y_2	d_1	d_2	D
1	-1	-1	-1	-1	-1	525	31,1	524,4	31,51	1,00000	0,00000	0,00000
2	1	-1	-1	-1	-1	531	34,0	531,0	33,86	0,99594	0,00000	0,00000
3	-1	1	-1	-1	-1	537	32,4	537,3	32,50	0,99150	0,00000	0,00000
4	1	1	-1	-1	-1	542	36,0	541,5	35,66	0,98843	0,00000	0,00000
5	-1	-1	1	-1	-1	595	35,9	596,1	35,61	0,93778	0,00000	0,00000
6	1	-1	1	-1	-1	602	40,1	601,8	40,02	0,93076	0,00000	0,00000
7	-1	1	1	-1	-1	603	37,7	602,3	37,46	0,93008	0,00000	0,00000
8	1	1	1	-1	-1	605	42,4	605,7	42,66	0,92571	0,00000	0,00000
9	-1	-1	-1	1	-1	555	32,3	555,0	32,44	0,97789	0,00000	0,00000
10	1	-1	-1	1	-1	561	35,8	561,8	35,49	0,97228	0,00000	0,00000
11	-1	1	-1	1	-1	567	34,1	566,8	33,88	0,96786	0,00000	0,00000
12	1	1	-1	1	-1	571	37,8	571,1	37,74	0,96391	0,00000	0,00000
13	-1	-1	1	1	-1	611	37,8	610,6	37,59	0,91889	0,00000	0,00000
14	1	-1	1	1	-1	617	42,6	616,5	42,70	0,91020	0,00000	0,00000
15	-1	1	1	1	-1	615	40,0	615,7	39,89	0,91130	0,00000	0,00000
16	1	1	1	1	-1	619	45,3	619,2	45,79	0,90583	0,83201	0,86814
17	-1	-1	-1	-1	1	592	37,9	591,4	37,46	0,94317	0,00000	0,00000
18	1	-1	-1	-1	1	597	42,4	597,2	42,39	0,93651	0,00000	0,00000
19	-1	1	-1	-1	1	604	40,0	604,7	39,88	0,92702	0,00000	0,00000
20	1	1	-1	-1	1	608	45,3	608,0	45,61	0,92252	0,81042	0,86465
21	-1	-1	1	-1	1	663	45,1	663,0	45,24	0,77668	0,73802	0,75711
22	1	-1	1	-1	1	668	52,0	667,9	52,22	0,73753	0,94300	0,83396
23	-1	1	1	-1	1	670	48,3	669,6	48,51	0,71661	0,96526	0,83169
24	1	1	1	-1	1	672	56,4	672,1	56,29	0,67380	0,00000	0,00000
25	-1	-1	-1	1	1	611	39,9	611,0	39,67	0,91839	0,00000	0,00000
26	1	-1	-1	1	1	617	45,0	616,8	45,30	0,90965	0,75421	0,82829
27	-1	1	-1	1	1	623	42,5	623,1	42,54	0,89931	0,00000	0,00000
28	1	1	-1	1	1	627	48,7	626,6	48,97	0,89312	0,97717	0,93420
29	-1	-1	1	1	1	666	48,3	666,4	48,50	0,75115	0,96490	0,85134
30	1	-1	1	1	1	671	56,2	671,4	56,18	0,68895	0,00000	0,00000
31	-1	1	1	1	1	672	51,9	671,9	52,22	0,67797	0,94302	0,79958
32	1	1	1	1	1	674	61,2	674,5	60,70	0,56406	0,00000	0,00000
33	-1	0	0	0	0	607	42,4	606,5	42,71	0,92464	0,00000	0,00000
34	1	0	0	0	0	612	48,5	611,1	48,12	0,91825	0,95408	0,93599
35	0	-1	0	0	0	606	44,1	605,8	44,34	0,92559	0,00000	0,00000
36	0	1	0	0	0	615	47,4	613,8	47,10	0,91429	0,91670	0,91549
37	0	0	-1	0	0	579	40,8	579,4	41,11	0,95599	0,00000	0,00000
38	0	0	1	0	0	641	49,4	639,2	49,02	0,86656	0,97853	0,92085
39	0	0	0	-1	0	607	44,3	607,0	44,43	0,92392	0,00000	0,00000
40	0	0	0	1	0	625	47,3	623,5	47,10	0,89854	0,91696	0,90770
41	0	0	0	0	-1	579	39,3	578,2	39,80	0,95716	0,00000	0,00000
42	0	0	0	0	1	640	50,8	639,4	50,23	0,86613	0,99529	0,92847
43	0	0	0	0	0	610	46,5	610,7	46,15	0,91881	0,86313	0,89053
44	0	0	0	0	0	607	46,1	610,7	46,15	0,91881	0,86313	0,89053
45	0	0	0	0	0	611	46,8	610,7	46,15	0,91881	0,86313	0,89053
46	0	0	0	0	0	609	46,5	610,7	46,15	0,91881	0,86313	0,89053
47	0	0	0	0	0	610	45,5	610,7	46,15	0,91881	0,86313	0,89053
48	0	0	0	0	0	608	45,8	610,7	46,15	0,91881	0,86313	0,89053
49	0	0	0	0	0	610	45,7	610,7	46,15	0,91881	0,86313	0,89053
50	0	0	0	0	0	612	46,2	610,7	46,15	0,91881	0,86313	0,89053
51	0	0	0	0	0	614	45,9	610,7	46,15	0,91881	0,86313	0,89053
52	0	0	0	0	0	610	46,2	610,7	46,15	0,91881	0,86313	0,89053

Table 5.16 Derringer-Suich optimization method results for *WMP*

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	525	525	675	0,1	1
Y2	Target	45	50	55	0,1	1

Global Solution

X1 = 1,00000
X2 = 1,00000
X3 = -0,13177
X4 = -0,81994
X5 = 0,41504

Predicted Responses

Y1 = 619,189; desirability = 0,90586
Y2 = 49,304; desirability = 0,98511

Composite Desirability = 0,94466

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	525	525	675	0,1	1
Y2	Target	45	50	55	1,0	1

Global Solution

X1 = 1,00000
X2 = 1,00000
X3 = -0,65316
X4 = -0,15499
X5 = 1,00000

Predicted Responses

Y1 = 622,118; desirability = 0,90099
Y2 = 49,951; desirability = 0,99027

Composite Desirability = 0,94458

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	525	525	675	0,1	1
Y2	Target	45	50	55	10,0	1

Global Solution

X1 = 1,00000
X2 = 1,00000
X3 = -0,56958
X4 = -0,41295
X5 = 1,00000

Predicted Responses

Y1 = 623,385; desirability = 0,89881
Y2 = 50,009; desirability = 0,98294

Composite Desirability = 0,93994

Table 5.16 ***(Continued)***

```
Parameters
                    Goal      Lower     Target     Upper     Weight     Import
Y1               Minimum        525        525       675        1,0          1
Y2                Target         45         50        55        0,1          1
Global Solution
X1            =  1,00000
X2            =  1,00000
X3            = -0,66751
X4            =  0,74528
X5            = -0,03172
Predicted Responses
Y1            =  603,848; desirability =  0,47434
Y2            =   46,345; desirability =  0,87693
Composite Desirability =  0,64495
```

```
Parameters
                    Goal      Lower     Target     Upper     Weight     Import
Y1               Minimum        525        525       675          1          1
Y2                Target         45         50        55          1          1
Global Solution
X1            =  1,00000
X2            =  1,00000
X3            = -0,65816
X4            = -0,12933
X5            =  1,00000
Predicted Responses
Y1            =  622,134; desirability =  0,35244
Y2            =   49,965; desirability =  0,99293
Composite Desirability =  0,59156
```

```
Parameters
                    Goal      Lower     Target     Upper     Weight     Import
Y1               Minimum        525        525       675          1          1
Y2                Target         45         50        55         10          1
Global Solution
X1            =  0,93946
X2            =  0,97824
X3            = -0,14255
X4            =  0,60680
X5            =  0,11530
Predicted Responses
Y1            =  619,817; desirability =  0,36788
Y2            =   49,999; desirability =  0,99821
Composite Desirability =  0,60599
```

Table 5.16 ***(Continued)***

```
Parameters
                 Goal      Lower     Target     Upper     Weight     Import
Y1            Minimum        525        525       675       10,0          1
Y2             Target         45         50        55        0,1          1
Global Solution
X1          =  1,00000
X2          =  0,97809
X3          = -0,56765
X4          = -0,23697
X5          = -0,13090
Predicted Responses
Y1          =  590,846; desirability =  0,00309
Y2          =   45,128; desirability =  0,69322
Composite Desirability =  0,04628
```

```
Parameters
                 Goal      Lower     Target     Upper     Weight     Import
Y1            Minimum        525        525       675         10          1
Y2             Target         45         50        55          1          1
Global Solution
X1          =  0,88775
X2          =  0,95907
X3          = -0,47202
X4          =  0,15499
X5          =  0,23145
Predicted Responses
Y1          =  608,624; desirability =  0,00029
Y2          =   48,139; desirability =  0,62790
Composite Desirability =  0,01344
```

```
Parameters
                 Goal      Lower     Target     Upper     Weight     Import
Y1            Minimum        525        525       675         10          1
Y2             Target         45         50        55         10          1
Global Solution
X1          =  1,00000
X2          =  1,00000
X3          = -0,00500
X4          =  0,13920
X5          =  0,08555
Predicted Responses
Y1          =  617,191; desirability =  0,00007
Y2          =   49,994; desirability =  0,98842
Composite Desirability =  0,00845
```

Since the optimum levels of input factors are found by using the metamodels, confirmatory simulation runs are needed at optimum input levels. The simulation model responses shown in Table 5.17 are the average values of 10 replications.

Table 5.17 Results of the confirmatory runs for *WMP*

	Weight		Composite	Natural variables					Metamodel		Simulation	
	Y_1	Y_2	desirability	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2	Y_1	Y_2
1	0,1	0,1	0,94466	600	600	430	327	512	619,2	49,30	622	49,4
2	0,1	1	0,94458	600	600	352	427	600	622,1	49,95	712	49,6
3	0,1	10	0,93994	600	600	365	388	600	623,4	50,01	628	50,1
4	1	0,1	0,64495	600	600	350	562	445	603,8	46,35	677	46,1
5	1	1	0,59156	600	600	351	431	600	622,1	49,97	711	49,7
6	1	10	0,60599	591	597	429	541	467	619,8	50,00	625	50,3
7	10	0,1	0,04628	600	597	365	414	430	590,8	45,13	595	45,4
8	10	1	0,01344	583	594	379	473	485	608,6	48,14	616	48,5
9	10	10	0,00845	600	600	449	471	463	617,2	49,99	625	50,3
10	Current		X	600	450	300	450	600	X	X	613	46,2

In this table, first nine rows reflect the global optimum solutions found by Derringer-Suich multi-response optimization method according to weight combinations, and the last row belongs to current situation.

Table 5.18 Factor values of confirmatory runs for *WMP*

	Coded Values					Natural values (second)					Natural values (minutes)				
	X_1	X_2	X_3	X_4	X_5	X_1	X_2	X_3	X_4	X_5	X_1	X_2	X_3	X_4	X_5
1	1,00	1,00	-0,13	-0,82	0,42	600	600	430	327	512	10,0	10,0	7,2	5,5	8,5
2	1,00	1,00	-0,65	-0,15	1,00	600	600	352	427	600	10,0	10,0	5,9	7,1	10,0
3	1,00	1,00	-0,57	-0,41	1,00	600	600	365	388	600	10,0	10,0	6,1	6,5	10,0
4	1,00	1,00	-0,67	0,75	-0,03	600	600	350	562	445	10,0	10,0	5,8	9,4	7,4
5	1,00	1,00	-0,66	-0,13	1,00	600	600	351	431	600	10,0	10,0	5,9	7,2	10,0
6	0,94	0,98	-0,14	0,61	0,12	591	597	429	541	467	9,8	9,9	7,1	9,0	7,8
7	1,00	0,98	-0,57	-0,24	-0,13	600	597	365	414	430	10,0	9,9	6,1	6,9	7,2
8	0,89	0,96	-0,47	0,15	0,23	583	594	379	473	485	9,7	9,9	6,3	7,9	8,1
9	1,00	1,00	-0,01	0,14	0,09	600	600	449	471	463	10,0	10,0	7,5	7,8	7,7
10	1,00	0,00	-1,00	0,00	1,00	600	450	300	450	600	10,0	7,5	5,0	7,5	10,0

In Table 5.18 Factor values of confirmatory runs for *WMP* are given in the first nine rows and the last row belongs to current situation.

First, the fullness rate of the carriages response (Y_2) is checked. In Table 5.17, we see that the fullness rate requirement 50% is provided by the second, third, fifth, sixth and ninth rows. After determining the natural factor levels providing the fullness rate requirement, the minimum value for the average passenger time spent in the metro-line (Y_1) is searched, and it is seen that the minimum Y_1 (625 seconds) is obtained with the natural factor levels in the 6th and 9th rows. As a result, two optimum points exist. For $X_1 = 9.8$, $X_2 = 9.9$, $X_3 = 7.1$, $X_4 = 9.0$, $X_5 = 7.8$ factor levels and $X_1 = 10.0$, $X_2 = 10.0$, $X_3 = 7.5$, $X_4 = 7.8$, $X_5 = 7.7$ factor levels the obtained values for the average passenger time spent in the metro-line (Y_1), and for the fullness rate of carriages (Y_2) responses are respectively 625 seconds and 50,3 percent. These factor levels are demonstrated in grey colour in Tables 5.17 and 5.18.

As can be seen from Table 5.17 after optimization study the fullness rate of carriages (Y_2) is increased by 4.1 percent and the average passenger time spent in the metro-line (Y_1) is increased by 12 seconds (1.96 percent) respect to current values. Although the current average passenger time spent in the metro-line (Y_1) value is less than optimum result, it violates the fullness rate of carriages requirement (50%).

5.2 Weekday Afternoon Problem (*WAP*)

5.2.1 Estimation Process

5.2.1.1 Phase Zero

The objective of *WAP* is to find the levels of 5 input factors that related to the 5 time periods from 11:30 a.m. to 00:00 a.m. and which minimize the average passenger time spent in the metro-line with the requirement as the fifty percent fullness rate of the carriages. The built simulation model is explained in chapter four.

Generally at phase zero, a screening experiment is made for investigating potential input factors, which are thought to be important in the response surface study, and for determining important factors. Because the input factors are 5 headways and none of them can be eliminated we skip factor screening processes.

The input factors are respectively;

X_1: The headway for the first time period from 11:30 a.m. to 17:00 p.m.

X_2: The headway for the second time period from 17:00 p.m. to 18:30 p.m.

X_3: The headway for the third time period from 18:30 p.m. to 19:00 p.m.

X_4: The headway for the fourth time period from 19:00 p.m. to 22:00 p.m.

X_5: The headway for the fifth time period from 22:00 p.m. to 00:0 a.m.

The output responses are respectively;

Y_1: The average passenger time that is spent in the metro-line (in second)

Y_2: The fullness rate of the carriages (in percentage)

5.2.1.2 Phase One

Low and high level of input factors for *WAP*, time period durations, and the natural values are given in Table 5.19.

Table 5.19 Low and high level of input factors for *WAP*

Input factor	Time Period	Duration		Low level		High level	
		(hour)	(second)	Coded	Natural (second)	Coded	Natural (second)
X_1	11:30 - 17:00	5,5	19800	-1	300	1	600
X_2	17:00 - 18:30	1,5	5400	-1	300	1	600
X_3	18:30 - 19:00	0,5	1800	-1	300	1	600
X_4	19:00 - 22:00	3,0	10800	-1	300	1	600
X_5	22:00 - 00:00	2,0	7200	-1	600	1	900

The first-order regression models with two-factor interactions are assumed to be as in equation (5.1).

5.2.1.2.1 Two-level Full Factorial Design

For fitting a first order regression model a two-level full factorial design (2^5) with central runs is designed and then the simulation model is run 10 times at each design point, and response values for each design points are found.

Although the values for two-level full factorial design points, which are denoted in Table 5.20, are average of 10 replications, the values related to the central points are obtained by one replication.

Table 5.20 Simulation results for *WAP* (2^5 design with 5 central runs)

	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2
1	-1	-1	-1	-1	-1	523	34,1
2	1	-1	-1	-1	-1	598	44,8
3	-1	1	-1	-1	-1	560	36,5
4	1	1	-1	-1	-1	629	48,9
5	-1	-1	1	-1	-1	545	34,9
6	1	-1	1	-1	-1	618	46,0
7	-1	1	1	-1	-1	569	37,4
8	1	1	1	-1	-1	638	50,5
9	-1	-1	-1	1	-1	555	39,2
10	1	-1	-1	1	-1	631	53,9
11	-1	1	-1	1	-1	592	42,4
12	1	1	-1	1	-1	663	60,1
13	-1	-1	1	1	-1	568	40,2
14	1	-1	1	1	-1	644	55,8
15	-1	1	1	1	-1	594	43,5
16	1	1	1	1	-1	663	62,4
17	-1	-1	-1	-1	1	531	35,1
18	1	-1	-1	-1	1	604	46,4
19	-1	1	-1	-1	1	566	37,6
20	1	1	-1	-1	1	637	50,8
21	-1	-1	1	-1	1	550	36,0
22	1	-1	1	-1	1	626	47,9
23	-1	1	1	-1	1	575	38,6
24	1	1	1	-1	1	645	52,8
25	-1	-1	-1	1	1	564	40,5
26	1	-1	-1	1	1	639	56,4
27	-1	1	-1	1	1	602	43,9
28	1	1	-1	1	1	669	63,1
29	-1	-1	1	1	1	576	41,7
30	1	-1	1	1	1	651	58,5
31	-1	1	1	1	1	602	45,2
32	1	1	1	1	1	672	65,9
33	0	0	0	0	0	609	50,6
34	0	0	0	0	0	612	50,4
35	0	0	0	0	0	608	50,1
36	0	0	0	0	0	610	50,2
37	0	0	0	0	0	609	50,1

After verifying that the residuals are normally, identically and independently distributed with zero mean and constant variance for the Y_1 and Y_2 responses, Variance Analysis is used to see if the main effects, two-factor interactions, curvature and lack-of-fit are statistically significant. Table 5.21 is the Analysis of Variance Table (ANOVA) for the first response Y_1, and Table 5.22 for the second response Y_2. In addition, estimated effects and coefficients (*EEC*) for two responses are given in Table 5.23 and 5.24, respectively.

Table 5.21 ANOVA table for Y_1 response for *form* for *WAP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	56411,4	56411,4	11282,3	1000,00	0,000
2-Way Interactions	10	457,3	457,3	45,7	4,70	0,001
Residual Error	21	204,3	204,3	9,7		
Curvature	1	183,1	183,1	183,1	172,69	0,000
Lack of Fit	16	12,0	12,0	0,8	0,33	0,953
Pure Error	4	9,2	9,2	2,3		
Total	36	57073,0				
S = 4,620 R-Sq = 98,8% R-Sq(adj) = 98,7%						

In Table 5.21, since the *P*-value for the lack-of-fit test is greater than the significance level (α = 0.05), the lack-of-fit is not statistically significant, that is, the first order regression model for Y_1 response for *WAP* is adequate. The main effects, two-way interactions and curvature are statistically significant (*P*-value<0.05).

Table 5.22 ANOVA table for Y_2 response for *form* for *WAP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	5	2516,2	2516,2	503,2	174,93	0,000
2-Way Interactions	10	79,3	79,3	7,9	2,76	0,024
Residual Error	21	60,4	60,4	2,9		
Curvature	1	58,8	58,8	58,8	710,96	0,000
Lack of Fit	16	1,5	1,5	0,1	1,95	0,273
Pure Error	4	0,2	0,2	0,0		
Total	36	2655,9				
S = 2,123 R-Sq = 94,7% R-Sq(adj) = 93,9%						

Table 5.22 shows that the lack-of-fit is not statically significant (P>0.05), curvature, main effects and the two-way interactions are statically significant (P<0.05).

As a result, the fitted first order models for both responses are respectively;

$$Y_1 = 603,973 + 36,094\, x_1 + 14,156\, x_2 + 5,406\, x_3 + 14,719\, x_4 + 3,719\, x_5 - 1,344\, x_1x_2 + 0,031\, x_1x_3 + 0,094\, x_1x_4 - 0,031\, x_1x_5 - 2,906\, x_2x_3 + 0,156\, x_2x_4 + 0,031\, x_2x_5 - 1,969\, x_3x_4 - 0,094\, x_3x_5 + 0,344\, x_4x_5$$

$$Y_2 = 47,092 + 7,419\, x_1 + 2,131\, x_2 + 0,738\, x_3 + 4,200\, x_4 + 0,931\, x_5 + 0,669\, x_1x_2 + 0,225\, x_1x_3 + 1,300\, x_1x_4 + 0,281\, x_1x_5 + 0,075\, x_2x_3 + 0,388\, x_2x_4 + 0,081\, x_2x_5 + 0,119\, x_3x_4 + 0,063\, x_3x_5 + 0,175\, x_4x_5 \quad (5.13)$$

Table 5.23 ***EEC*** **table for** Y_1 **response for** ***form*** **for** ***WAP***

Term	Effect	Coef	SE Coef	T	P
Constant		603,973	0,5127	1177,99	0,000
X1	72,188	36,094	0,5513	65,47	0,000
X2	28,312	14,156	0,5513	25,68	0,000
X3	10,812	5,406	0,5513	9,81	0,000
X4	29,437	14,719	0,5513	26,70	0,000
X5	7,438	3,719	0,5513	6,75	0,000
*X1*X2*	-2,687	-1,344	0,5513	-2,44	0,024
*X1*X3*	0,062	0,031	0,5513	0,06	0,955
*X1*X4*	0,188	0,094	0,5513	0,17	0,867
*X1*X5*	-0,062	-0,031	0,5513	-0,06	0,955
*X2*X3*	-5,812	-2,906	0,5513	-5,27	0,000
*X2*X4*	0,313	0,156	0,5513	0,28	0,780
*X2*X5*	0,063	0,031	0,5513	0,06	0,955
*X3*X4*	-3,937	-1,969	0,5513	-3,57	0,002
*X3*X5*	-0,188	-0,094	0,5513	-0,17	0,867
*X4*X5*	0,688	0,344	0,5513	0,62	0,540

Table 5.24 ***EEC*** **table for** Y_2 **response for** ***form*** **for** ***WAP***

Term	Effect	Coef	SE Coef	T	P
Constant		47,092	0,2788	168,88	0,000
X1	14,838	7,419	0,2998	24,74	0,000
X2	4,263	2,131	0,2998	7,11	0,000
X3	1,475	0,738	0,2998	2,46	0,023
X4	8,400	4,200	0,2998	14,01	0,000
X5	1,863	0,931	0,2998	3,11	0,005
*X1*X2*	1,338	0,669	0,2998	2,23	0,037
*X1*X3*	0,450	0,225	0,2998	0,75	0,461
*X1*X4*	2,600	1,300	0,2998	4,34	0,000
*X1*X5*	0,563	0,281	0,2998	0,94	0,359
*X2*X3*	0,150	0,075	0,2998	0,25	0,805
*X2*X4*	0,775	0,388	0,2998	1,29	0,210
*X2*X5*	0,163	0,081	0,2998	0,27	0,789
*X3*X4*	0,238	0,119	0,2998	0,40	0,696
*X3*X5*	0,125	0,063	0,2998	0,21	0,837
*X4*X5*	0,350	0,175	0,2998	0,58	0,566

Because the true response surface usually exhibits curvature near the optimum, it is understood that the determined region of experimentation is near the region of optimum. As mentioned in chapter two, the second phase of *RSM* begins. Therefore, second order models can be built.

5.2.2 Optimization Process

5.2.2.1 Phase Two

5.2.2.1.1 *CCD* and Development of Metamodels for *WAP*

The second order regression models are assumed to be as in equation (5.10).

For fitting second order models a *CCD* is built by augmenting the two-level full factorial design with central runs and axial runs ($\alpha = 1$, means design is face-centered). In Table 5.25, the *CCD* used for fitting second order models for two responses is given. The values for two-level full factorial design points and axial points are average of 10 replications, and the values related to the central points are obtained by one replication.

After verifying that the residuals of the second order regression models fitted for Y_1 and Y_2 are normally, identically and independently distributed with zero mean and constant variance, the ANOVA tables are given in Table 5.26 and Table 5.27. Also estimated effects and coefficients (*EEC*) for two responses are denoted in Table 5.28 and 5.29 respectively.

Table 5.25 Simulation results for *WAP* (*CCD*)

	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2		X_1	X_2	X_3	X_4	X_5	Y_1	Y_2
1	-1	-1	-1	-1	-1	523	34,1	27	-1	1	-1	1	1	602	43,9
2	1	-1	-1	-1	-1	598	44,8	28	1	1	-1	1	1	669	63,1
3	-1	1	-1	-1	-1	560	36,5	29	-1	-1	1	1	1	576	41,7
4	1	1	-1	-1	-1	629	48,9	30	1	-1	1	1	1	651	58,5
5	-1	-1	1	-1	-1	545	34,9	31	-1	1	1	1	1	602	45,2
6	1	-1	1	-1	-1	618	46,0	32	1	1	1	1	1	672	65,9
7	-1	1	1	-1	-1	569	37,4	33	-1	0	0	0	0	567	40,7
8	1	1	1	-1	-1	638	50,5	34	1	0	0	0	0	640	56,7
9	-1	-1	-1	1	-1	555	39,2	35	0	-1	0	0	0	595	47,2
10	1	-1	-1	1	-1	631	53,9	36	0	1	0	0	0	625	51,8
11	-1	1	-1	1	-1	592	42,4	37	0	0	-1	0	0	607	49,1
12	1	1	-1	1	-1	663	60,1	38	0	0	1	0	0	621	50,6
13	-1	-1	1	1	-1	568	40,2	39	0	0	0	-1	0	593	44,5
14	1	-1	1	1	-1	644	55,8	40	0	0	0	1	0	623	53,6
15	-1	1	1	1	-1	594	43,5	41	0	0	0	0	-1	606	48,9
16	1	1	1	1	-1	663	62,4	42	0	0	0	0	1	614	50,9
17	-1	-1	-1	-1	1	531	35,1	43	0	0	0	0	0	610	50,0
18	1	-1	-1	-1	1	604	46,4	44	0	0	0	0	0	610	50,5
19	-1	1	-1	-1	1	566	37,6	45	0	0	0	0	0	608	49,9
20	1	1	-1	-1	1	637	50,8	46	0	0	0	0	0	608	49,9
21	-1	-1	1	-1	1	550	36,0	47	0	0	0	0	0	610	50,0
22	1	-1	1	-1	1	626	47,9	48	0	0	0	0	0	608	50,1
23	-1	1	1	-1	1	575	38,6	49	0	0	0	0	0	612	50,5
24	1	1	1	-1	1	645	52,8	50	0	0	0	0	0	610	50,4
25	-1	-1	-1	1	1	564	40,5	51	0	0	0	0	0	608	50,2
26	1	-1	-1	1	1	639	56,4	52	0	0	0	0	0	611	50,0

Table 5.26 ANOVA table for Y_1 response for *sorm* for *WAP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	20	61143,1	61143,1	3057,2	2000,00	0,000
Linear	5	60099,2	60099,2	12019,8	9000,00	0,000
Square	5	586,6	586,6	117,3	86,73	0,000
Interaction	10	457,3	457,3	45,7	33,81	0,000
Residual Error	31	41,9	41,9	1,4		
Lack-of-Fit	22	23,4	23,4	1,1	0,52	0,900
Pure Error	9	18,5	18,5	2,1		
Total	51	61185,0				
S = 1,163 R-Sq = 99,9% R-Sq(adj) = 99,9%						

Table 5.27 ANOVA table for Y_2 response for *sorm* for *WAP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	20	2907,2	2907,2	145,4	2000,00	0,000
Linear	5	2698,4	2698,4	539,7	6000,00	0,000
Square	5	129,5	129,5	25,9	276,08	0,000
Interaction	10	79,3	79,3	7,9	84,52	0,000
Residual Error	31	2,9	2,9	0,1		
Lack-of-Fit	22	2,4	2,4	0,1	1,95	0,151
Pure Error	9	0,5	0,5	0,1		
Total	51	2910,1				
S = 0,3063 R-Sq = 99,9% R-Sq(adj) = 99,8%						

In Table 5.26 and Table 5.27, we see that linear, quadratic and two-way interaction effects are statistically significant (P<0.05), and the second order regression models are statistically significant (P>0.05) for both responses.

Table 5.28 *EEC* table for Y_1 response for *sorm* for *WAP*

Term	Coef	SE Coef	T	P
Constant	609,927	0,288	2118,7160	0,000
X1	36,118	0,200	181,0840	0,000
X2	14,206	0,200	71,2250	0,000
X3	5,500	0,200	27,5760	0,000
X4	14,735	0,200	73,8790	0,000
X5	3,735	0,200	18,7280	0,000
*X1*X1*	-6,960	0,739	-9,4160	0,000
*X2*X2*	-0,460	0,739	-0,6220	0,538
*X3*X3*	3,540	0,739	4,7890	0,000
*X4*X4*	-2,460	0,739	-3,3280	0,002
*X5*X5*	-0,460	0,739	-0,6220	0,538
*X1*X2*	-1,344	0,206	-6,5360	0,000
*X1*X3*	0,031	0,206	0,1520	0,880
*X1*X4*	0,094	0,206	0,4560	0,652
*X1*X5*	-0,031	0,206	-0,1520	0,880
*X2*X3*	-2,906	0,206	-14,1360	0,000
*X2*X4*	0,156	0,206	0,7600	0,453
*X2*X5*	0,031	0,206	0,1520	0,880
*X3*X4*	-1,969	0,206	-9,5760	0,000
*X3*X5*	-0,094	0,206	-0,4560	0,652
*X4*X5*	0,344	0,206	1,6720	0,105

Table 5.29 *EEC* table for Y_2 response for *sorm* for *WAP*

Term	Coef	SE Coef	T	P
Constant	50,131	0,076	661,2010	0,000
X1	7,453	0,053	141,8790	0,000
X2	2,141	0,053	40,7610	0,000
X3	0,738	0,053	14,0540	0,000
X4	4,221	0,053	80,3460	0,000
X5	0,935	0,053	17,8050	0,000
*X1*X1*	-1,408	0,195	-7,2320	0,000
*X2*X2*	-0,608	0,195	-3,1220	0,004
*X3*X3*	-0,258	0,195	-1,3240	0,195
*X4*X4*	-1,058	0,195	-5,4340	0,000
*X5*X5*	-0,208	0,195	-1,0670	0,294
*X1*X2*	0,669	0,054	12,3510	0,000
*X1*X3*	0,225	0,054	4,1550	0,000
*X1*X4*	1,300	0,054	24,0090	0,000
*X1*X5*	0,281	0,054	5,1940	0,000
*X2*X3*	0,075	0,054	1,3850	0,176
*X2*X4*	0,388	0,054	7,1560	0,000
*X2*X5*	0,081	0,054	1,5010	0,144
*X3*X4*	0,119	0,054	2,1930	0,036
*X3*X5*	0,063	0,054	1,1540	0,257
*X4*X5*	0,175	0,054	3,2320	0,003

As a result, the second order regression models (metamodels) for Y_1 and Y_2 are respectively;

$$Y_1 = 609{,}927 + 36{,}118\, x_1 + 14{,}206\, x_2 + 5{,}500\, x_3 + 14{,}735\, x_4 + 3{,}735\, x_5 - 6{,}960\, x_1^2 - 0{,}460\, x_2^2 + 3{,}540\, x_3^2 - 2{,}460\, x_4^2 - 0{,}460\, x_5^2 - 1{,}344\, x_1x_2 + 0{,}031\, x_1x_3 + 0{,}094\, x_1x_4 - 0{,}031\, x_1x_5 - 2{,}906\, x_2x_3 + 0{,}156\, x_2x_4 + 0{,}031\, x_2x_5 - 1{,}969\, x_3x_4 - 0{,}094\, x_3x_5 + 0{,}344\, x_4x_5$$

$$Y_2 = 50{,}131 + 7{,}453\, x_1 + 2{,}141\, x_2 + 0{,}738\, x_3 + 4{,}221\, x_4 + 0{,}935\, x_5 - 1{,}408\, x_1^2 - 0{,}608\, x_2^2 - 0{,}258\, x_3^2 - 1{,}058\, x_4^2 - 0{,}208\, x_5^2 + 0{,}669\, x_1x_2 + 0{,}225\, x_1x_3 + 1{,}300\, x_1x_4 + 0{,}281\, x_1x_5 + 0{,}075\, x_2x_3 + 0{,}388\, x_2x_4 + 0{,}081\, x_2x_5 + 0{,}119\, x_3x_4 + 0{,}063\, x_3x_5 + 0{,}175\, x_4x_5 \qquad (5.14)$$

After building regression metamodels for both responses, they are verified and validated.

5.2.2.1.2 Verification and Validation of *WAP* Metamodels

For verification, lack-of-fit tests are applied for both metamodels. As can be seen from Table 5.26 and Table 5.27, the metamodels have no statistically significant lack-of-fit (P>0.05) for both responses with %5 significance level ($\alpha = 0.05$). R^2 is 99.9%, and R^2-*adjusted* is 99.9% for response Y_1, and R^2 is 99.9%, and R^2-*adjusted* is 99.8% for response Y_2.

For validation, 32 (2^5) points (because there are 5 input factors with 2 level and we want to see if metamodel is valid in the entire experimentation region) that selected randomly in experimentation region, and from different design points, are used. These 32 points are shown in Table 5.30. The simulation model is run at these randomly selected points with 10 replications for each point. The values that are denoted in the Simulation Model column of Table 5.30 are the average these 10 replications. Also the values denoted in Metamodel column are predicted by fitted second order regression metamodels. Since the entire *ARE* values are smaller than 3% the second order metamodels can be used for prediction.

After verification and validation of regression metamodels, Derringer-Suich multi-response optimization procedure is used for optimization.

5.2.2.1.3 Derringer-Suich Multi-Response Optimization Procedure for *WAP*

In the current problem, since the target value for the fullness rate of the carriages response (Y_2) is 50%, the lower and upper limits are determined as 45% and 55%, respectively. To decide the upper and lower limits on the average passenger time spent in the metro-line response (Y_1) the simulation model is run 20 times at the low levels of all input factors. The lower limit is defined as the average of these 20 values, and is found as 523 seconds. To determine the upper limit, the simulation model is run 20 times at the high levels of all input factors, and the average value, 672 seconds, is found as upper limit.

Table 5.30 Metamodels validation for *WAP*

	Input factors					Simulation model		Metamodel		ARE (%)	
	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
1	-0,6	-0,2	-0,6	-0,2	-0,6	576	42,9	574,7	43,15	0,23	0,59
2	0,2	-0,2	-0,2	-0,6	-0,6	611	47,4	600,7	47,32	1,68	0,16
3	-0,6	0,2	-0,6	-0,2	-0,2	581	44,0	583,0	44,13	0,35	0,30
4	0,6	0,6	-0,2	-0,6	-0,2	623	51,4	625,6	51,49	0,41	0,17
5	-0,2	-0,2	0,6	-0,6	-0,2	594	45,3	594,7	45,55	0,12	0,56
6	0,2	-0,6	0,6	-0,2	-0,6	609	48,8	608,9	48,78	0,02	0,03
7	-0,2	0,6	0,2	-0,6	-0,2	596	46,6	601,6	46,62	0,94	0,05
8	0,2	0,2	0,2	-0,2	-0,6	620	50,5	615,4	50,52	0,74	0,04
9	-0,6	-0,6	-0,2	0,6	-0,6	581	44,7	580,8	44,74	0,03	0,08
10	0,6	-0,6	-0,6	0,2	-0,2	616	52,5	620,1	52,44	0,67	0,11
11	-0,2	0,2	-0,6	0,6	-0,2	609	50,3	611,5	50,24	0,41	0,11
12	0,2	0,2	-0,2	0,6	-0,6	631	53,5	624,5	53,43	1,03	0,13
13	-0,2	-0,2	0,2	0,6	-0,6	609	49,7	606,0	49,58	0,49	0,23
14	0,2	-0,2	0,6	0,2	-0,2	624	52,2	620,8	52,09	0,51	0,21
15	-0,6	0,6	0,6	0,2	-0,2	595	46,6	600,0	46,79	0,83	0,42
16	0,6	0,6	0,6	0,2	-0,6	642	56,1	640,7	56,07	0,20	0,05
17	-0,6	-0,6	-0,2	-0,6	0,2	563	41,4	566,1	41,59	0,55	0,45
18	0,6	-0,2	-0,6	-0,6	0,2	615	50,0	614,3	49,84	0,12	0,31
19	-0,2	0,6	-0,6	-0,2	0,6	605	48,8	608,8	48,61	0,62	0,40
20	0,2	0,2	-0,2	-0,6	0,6	622	49,1	610,8	49,18	1,81	0,17
21	-0,6	-0,2	0,6	-0,2	0,2	581	44,2	585,6	44,45	0,78	0,57
22	0,6	-0,6	0,2	-0,2	0,2	620	52,0	620,2	51,70	0,04	0,57
23	-0,2	0,6	0,6	-0,6	0,6	603	47,5	607,3	47,42	0,72	0,16
24	0,6	0,6	0,2	-0,2	0,2	635	54,9	635,6	54,70	0,09	0,36
25	-0,6	-0,2	-0,6	0,6	0,2	586	45,9	589,8	46,08	0,64	0,40
26	0,6	-0,2	-0,2	0,2	0,6	627	55,2	630,4	54,91	0,54	0,52
27	-0,2	0,2	-0,6	0,6	0,6	613	51,1	614,5	50,95	0,25	0,30
28	0,2	0,2	-0,2	0,6	0,2	637	54,3	627,8	54,38	1,45	0,14
29	-0,2	-0,6	0,6	0,2	0,2	597	48,2	602,5	48,35	0,92	0,30
30	0,2	-0,6	0,6	0,2	0,6	621	51,6	618,6	51,69	0,38	0,17
31	-0,6	0,6	0,2	0,2	0,2	596	46,8	599,0	46,96	0,50	0,34
32	0,6	0,2	0,2	0,2	0,2	638	56,0	636,4	55,95	0,25	0,09

We will use three values 0.1, 1 and 10 for t and s. Therefore, nine situations (combination of weights) appear to be evaluated.

For each combination of weights, the lower, target, upper values and weights for two responses, and the optimum conditions with coded variables, and also corresponding predicted responses at these coded values which are calculated from the fitted models, in addition the individual desirability values and the composite desirability values at optimum conditions are shown in Table 5.31.

Table 5.31 Derringer-Suich optimization method results for *WAP*

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	523	523	672	0,1	1
Y2	Target	45	50	55	0,1	1

Global Solution

X1 = 0,00000
X2 = 0,00000
X3 = 0,00000
X4 = -0,01224
X5 = -0,08355

Predicted Responses

Y1 = 609,431; desirability = 0,91689
Y2 = 50,000; desirability = 1,00000

Composite Desirability = 0,95754

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	523	523	672	0,1	1
Y2	Target	45	50	55	1,0	1

Global Solution

X1 = 0,01283
X2 = -0,66928
X3 = -0,21012
X4 = 1,00000
X5 = -1,00000

Predicted Responses

Y1 = 607,327; desirability = 0,91993
Y2 = 50,000; desirability = 1,00000

Composite Desirability = 0,95913

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	523	523	672	0,1	1
Y2	Target	45	50	55	10,0	1

Global Solution

X1 = 0,15576
X2 = -0,94380
X3 = -0,18391
X4 = 0,91108
X5 = -1,00000

Predicted Responses

Y1 = 607,468; desirability = 0,91973
Y2 = 50,000; desirability = 1,00000

Composite Desirability = 0,95902

Table 5.31 ***(Continued)***

```
Parameters
                    Goal      Lower    Target     Upper    Weight    Import
Y1               Minimum        523       523       672       1,0         1
Y2                Target         45        50        55       0,1         1
Global Solution
X1            = -0,27707
X2            = -1,00000
X3            = -0,35656
X4            =  0,31770
X5            =  1,00000
Predicted Responses
Y1            =  589,796; desirability =  0,55170
Y2            =   46,732; desirability =  0,89942
Composite Desirability =  0,70442
```

```
Parameters
                    Goal      Lower    Target     Upper    Weight    Import
Y1               Minimum        523       523       672         1         1
Y2                Target         45        50        55         1         1
Global Solution
X1            =  0,17949
X2            = -1,00000
X3            = -0,32685
X4            =  1,00000
X5            = -1,00000
Predicted Responses
Y1            =  607,636; desirability =  0,43197
Y2            =   50,000; desirability =  1,00000
Composite Desirability =  0,65725
```

```
Parameters
                    Goal      Lower    Target     Upper    Weight    Import
Y1               Minimum        523       523       672         1         1
Y2                Target         45        50        55        10         1
Global Solution
X1            =  0,15576
X2            = -0,94380
X3            = -0,22805
X4            =  0,93086
X5            = -1,00000
Predicted Responses
Y1            =  607,444; desirability =  0,43326
Y2            =   50,000; desirability =  1,00000
Composite Desirability =  0,65823
```

Table 5.31 ***(Continued)***

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	523	523	672	10,0	1
Y2	Target	45	50	55	0,1	1

Global Solution

X1 = -0,30829
X2 = -0,86116
X3 = -0,36417
X4 = -0,00083
X5 = -0,10163

Predicted Responses

Y1 = 582,359; desirability = 0,00621
Y2 = 45,247; desirability = 0,74013

Composite Desirability = 0,06780

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	523	523	672	10	1
Y2	Target	45	50	55	1	1

Global Solution

X1 = -0,27707
X2 = -1,00000
X3 = -0,56048
X4 = 0,37586
X5 = 1,00000

Predicted Responses

Y1 = 589,724; desirability = 0,00264
Y2 = 46,707; desirability = 0,34133

Composite Desirability = 0,02999

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	523	523	672	10	1
Y2	Target	45	50	55	10	1

Global Solution

X1 = -0,07964
X2 = -0,50059
X3 = -0,29753
X4 = 1,00000
X5 = -0,78503

Predicted Responses

Y1 = 607,249; desirability = 0,00024
Y2 = 50,000; desirability = 0,99931

Composite Desirability = 0,01549

Since the optimum levels of input factors are found by using the metamodels, confirmatory simulation runs are needed at optimum input levels. The simulation model responses shown in Table 5.32 are the average values of 10 replications.

Table 5.32 Results of the confirmatory runs for *WAP*

	Weight		Composite	Natural variables					Metamodel		Simulation	
	Y_1	Y_2	desirability	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2	Y_1	Y_2
1	0,1	0,1	0,95754	450	450	450	448	737	609,4	50,00	608	49,9
2	0,1	1	0,95913	452	350	418	600	600	607,3	50,00	667	50,1
3	0,1	10	0,95902	473	308	422	587	600	607,5	50,00	612	49,9
4	1	0,1	0,70442	408	300	397	498	900	589,8	46,73	628	46,4
5	1	1	0,65725	477	300	401	600	600	607,6	50,00	614	50,0
6	1	10	0,65823	473	308	416	590	600	607,4	50,00	612	49,9
7	10	0,1	0,06780	404	321	395	450	735	582,4	45,25	577	45,1
8	10	1	0,02999	408	300	366	506	900	589,7	46,71	620	46,5
9	10	10	0,01549	438	375	405	600	632	607,2	50,00	606	50,2
10	Current		X	450	300	450	600	900	X	X	613	50,9

In this table, first nine rows reflect the global optimum solutions found by Derringer-Suich multi-response optimization method according to weight combinations, and the last row belongs to current situation.

Table 5.33 Factor values of confirmatory runs for *WAP*

	Coded values					Natural values (second)					Natural values (minutes)				
	X_1	X_2	X_3	X_4	X_5	X_1	X_2	X_3	X_4	X_5	X_1	X_2	X_3	X_4	X_5
1	0,00	0,00	0,00	-0,01	-0,08	450	450	450	448	737	7,5	7,5	7,5	7,5	12,3
2	0,01	-0,67	-0,21	1,00	-1,00	452	350	418	600	600	7,5	5,8	7,0	10,0	10,0
3	0,16	-0,94	-0,18	0,91	-1,00	473	308	422	587	600	7,9	5,1	7,0	9,8	10,0
4	-0,28	-1,00	-0,36	0,32	1,00	408	300	397	498	900	6,8	5,0	6,6	8,3	15,0
5	0,18	-1,00	-0,33	1,00	-1,00	477	300	401	600	600	7,9	5,0	6,7	10,0	10,0
6	0,16	-0,94	-0,23	0,93	-1,00	473	308	416	590	600	7,9	5,1	6,9	9,8	10,0
7	-0,31	-0,86	-0,36	0,00	-0,10	404	321	395	450	735	6,7	5,3	6,6	7,5	12,2
8	-0,28	-1,00	-0,56	0,38	1,00	408	300	366	506	900	6,8	5,0	6,1	8,4	15,0
9	-0,08	-0,50	-0,30	1,00	-0,79	438	375	405	600	632	7,3	6,2	6,8	10,0	10,5
10	0,00	-1,00	0,00	1,00	1,00	450	300	450	600	900	7,5	5,0	7,5	10,0	15,0

In Table 5.33, factor values of confirmatory runs for *WAP* are given in the first nine rows and the last row belongs to current situation.

First, the fullness rate of the carriages response (Y_2) is checked. In Table 5.32, we see that the fullness rate requirement 50% is provided by the first, second, third, fifth, sixth and ninth rows. After determining the natural factor levels providing the fullness rate requirement, the minimum value for the average passenger time spent in the metro-line (Y_1) is searched, and it is seen that the minimum Y_1 (606 seconds) is obtained with the natural factor levels in the 9th row. As a result, one optimum point exist. For $X_1 = 7.3, X_2 = 6.2, X_3 = 6.8, X_4 = 10.0, X_5 = 10.5$ factor levels the obtained values for the average passenger time spent in the metro-line (Y_1), and for the fullness rate of carriages (Y_2) responses are respectively 606 seconds and 50,2 percent. This optimum point's factor levels are demonstrated in grey colour in Tables 5.32 and 5.33.

As can be seen from Table 5.32 after optimization study the fullness rate of carriages (Y_2) is decreased by 0.7 percent and the average passenger time spent in the metro-line (Y_1) is decreased by 7 seconds (1.14 percent) respect to current values.

5.3 Saturday Problem (*STP*)

5.3.1 Estimation Process

5.3.1.1 Phase Zero

The objective of *STP* is to find the levels of 4 input factors that related to the 4 time periods from 06:00 a.m. to 00:00 a.m. and which minimize the average passenger time spent in the metro-line with the requirement as the fifty percent fullness rate of the carriages. The built simulation model is explained in chapter four.

Generally at phase zero, a screening experiment is made for investigating potential input factors, which are thought to be important in the response surface study, and for determining important factors. Because the input factors are 4 headways and none of them can be eliminated we skip factor screening processes.

The input factors are respectively;

X_1: The headway for the first time period from 06:00 a.m. to 11:00 a.m.

X_2: The headway for the second time period from 11:00 a.m. to 19:00 p.m.

X_3: The headway for the third time period from 19:00 p.m. to 22:00 p.m.

X_4: The headway for the fourth time period from 22:00 p.m. to 00:00 a.m.

The output responses are respectively;

Y_1: The average passenger time that is spent in the metro-line (in second)

Y_2: The fullness rate of the carriages (in percentage)

5.3.1.2 Phase One

Low and high level of input factors for *STP*, time period durations, and the natural values are given in Table 5.34.

Table 5.34 Low and high level of input factors for *STP*

Input factor	Time Period	Duration		Low level		High level	
		(hour)	(second)	Coded	Natural (second)	Coded	Natural (second)
X_1	06:00 - 11:00	5,0	18000	-1	300	1	600
X_2	11:00 - 19:00	8,0	28800	-1	300	1	600
X_3	19:00 - 22:00	3,0	10800	-1	300	1	600
X_4	22:00 - 00:00	2,0	7200	-1	600	1	900

The first order regression models with two-factor interactions are assumed to be;

$$Y_1 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{14} X_1 X_4 + \beta_{23} X_2 X_3 + \beta_{24} X_2 X_4 + \beta_{34} X_3 X_4 + \varepsilon$$

$$Y_2 = \delta_0 + \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3 + \delta_4 X_4 + \delta_{12} X_1 X_2 + \delta_{13} X_1 X_3 + \delta_{14} X_1 X_4 + \delta_{23} X_2 X_3 + \delta_{24} X_2 X_4 + \delta_{34} X_3 X_4 + \varepsilon \qquad (5.15)$$

where β_0 and δ_0 are constant, and β_1, β_2, β_3, β_4, δ_1, δ_2, δ_3 and δ_4 are coefficients corresponding to main effects, and β_{12}, β_{13}, β_{14}, β_{23}, β_{24}, β_{34}, δ_{12}, δ_{13}, δ_{14}, δ_{23}, δ_{24}, and δ_{34} are coefficients corresponding to two-factor interaction effects and ε is statistical error that have a normal distribution with mean zero and variance σ^2.

5.3.1.2.1 Two-level Full Factorial Design

For fitting a first order regression model a two-level full factorial design (2^4) with central runs is designed and then the simulation model is run 10 times at each design point, and response values for each design points are found.

Although the values for two-level full factorial design points, which are denoted in Table 5.35, are average of 10 replications, the values related to the central points are obtained by one replication.

Table 5.35 Simulation results for *STP* (2^4 design with 5 central runs)

	X_1	X_2	X_3	X_4	Y_1	Y_2
1	-1	-1	-1	-1	523	35,3
2	1	-1	-1	-1	555	41,4
3	-1	1	-1	-1	621	46,1
4	1	1	-1	-1	653	57,2
5	-1	-1	1	-1	540	38,7
6	1	-1	1	-1	573	46,0
7	-1	1	1	-1	634	52,0
8	1	1	1	-1	665	66,4
9	-1	-1	-1	1	526	36,0
10	1	-1	-1	1	561	42,3
11	-1	1	-1	1	625	47,2

	X_1	X_2	X_3	X_4	Y_1	Y_2
12	1	1	-1	1	657	58,9
13	-1	-1	1	1	545	39,6
14	1	-1	1	1	579	47,2
15	-1	1	1	1	640	53,6
16	1	1	1	1	671	68,9
17	0	0	0	0	605	52,3
18	0	0	0	0	609	51,8
19	0	0	0	0	606	52,4
20	0	0	0	0	606	52,5
21	0	0	0	0	609	52,6

After verifying that the residuals are normally, identically and independently distributed with zero mean and constant variance for the Y_1 and Y_2 responses, Variance Analysis is used to see if the main effects, two-factor interactions, curvature and lack-of-fit are statistically significant. Table 5.36 is the Analysis of Variance Table (ANOVA) for the first response Y_1, and Table 5.37 for the second response Y_2. In addition, estimated effects and coefficients (*EEC*) for two responses are given in Table 5.38 and 5.39, respectively.

Table 5.36 ANOVA table for Y_1 response for *form* for *STP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	41798,2	41798,2	10449,6	321,95	0,000
2-Way Interactions	6	27,7	27,7	4,6	0,14	0,987
Residual Error	10	324,6	324,6	32,5		
Curvature	1	308,6	308,6	308,6	173,57	0,000
Lack of Fit	5	2,0	2,0	0,4	0,11	0,982
Pure Error	4	14,0	14,0	3,5		
Total	20	42150,6				
S = 4,693 R-Sq = 99,2% R-Sq(adj) = 99,0%						

In Table 5.36, since the P-value for the lack-of-fit test is greater than the significance level ($\alpha = 0.05$), the lack-of-fit is not statistically significant, that is, the first order regression model for Y_1 response for *WAP* is adequate. The main effects and curvature are statistically significant (P-value<0.05), two-way interactions are not statistically significant (P-value>0.05).

Table 5.37 ANOVA table for Y_2 response for *form* for *STP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	1506,9	1506,9	376,7	67,45	0,000
2-Way Interactions	6	60,4	60,4	10,1	1,80	0,196
Residual Error	10	55,9	55,9	5,6		
Curvature	1	54,1	54,1	54,1	284,47	0,000
Lack of Fit	5	1,3	1,3	0,3	2,73	0,176
Pure Error	4	0,4	0,4	0,1		
Total	20	1623,2				
S = 2,695 R-Sq = 92,8% R-Sq(adj) = 91,0%						

Table 5.37 shows that the lack-of-fit and two-way interactions are not statically significant (P>0.05), curvature and main effects are statically significant (P<0.05).

Table 5.38 *EEC* table for Y_1 response for *form* for *STP*

Term	Effect	Coef	SE Coef	T	P
Constant		600,143	1,2430	482,74	0,000
X1	32,500	16,250	1,4240	11,41	0,000
X2	95,500	47,750	1,4240	33,53	0,000
X3	15,750	7,875	1,4240	5,53	0,000
X4	5,000	2,500	1,4240	1,76	0,110
*X1*X2*	-1,000	-0,500	1,4240	-0,35	0,733
*X1*X3*	-0,250	-0,125	1,4240	-0,09	0,932
*X1*X4*	0,500	0,250	1,4240	0,18	0,864
*X2*X3*	-2,250	-1,125	1,4240	-0,79	0,448
*X2*X4*	0,000	0,000	1,4240	0,00	1,000
*X3*X4*	0,750	0,375	1,4240	0,26	0,798

Table 5.39 *EEC* table for Y_2 response for *form* for *STP*

Term	Effect	Coef	SE Coef	T	P
Constant		49,448	0,5157	95,88	0,000
X1	9,975	4,988	0,5909	8,44	0,000
X2	15,475	7,738	0,5909	13,10	0,000
X3	6,000	3,000	0,5909	5,08	0,000
X4	1,325	0,663	0,5909	1,12	0,288
*X1*X2*	3,150	1,575	0,5909	2,67	0,024
*X1*X3*	1,175	0,588	0,5909	0,99	0,344
*X1*X4*	0,250	0,125	0,5909	0,21	0,837
*X2*X3*	1,875	0,938	0,5909	1,59	0,144
*X2*X4*	0,400	0,200	0,5909	0,34	0,742
*X3*X4*	0,225	0,113	0,5909	0,19	0,853

As a result, the fitted first order models for both responses are respectively;

$$Y_1 = 600{,}143 + 16{,}250\, x_1 + 47{,}750\, x_2 + 7{,}875\, x_3 + 2{,}500\, x_4 - 0{,}500\, x_1x_2 - 0{,}125\, x_1x_3 + 0{,}250\, x_1x_4 - 1{,}125\, x_2x_3 + 0{,}375\, x_3x_4$$

$$Y_2 = 49{,}448 + 4{,}988\, x_1 + 7{,}738\, x_2 + 3{,}000\, x_3 + 0{,}663\, x_4 + 1{,}575\, x_1x_2 + 0{,}588\, x_1x_3 + 0{,}125\, x_1x_4 + 0{,}938\, x_2x_3 + 0{,}200\, x_2x_4 + 0{,}113\, x_3x_4 \quad (5.16)$$

Because the true response surface usually exhibits curvature near the optimum, it is understood that the determined region of experimentation is near the region of optimum. As mentioned in chapter two, the second phase of *RSM* begins. Therefore, second order models can be built.

5.3.2 Optimization Process

5.3.2.1 Phase Two

5.3.2.1.1 *CCD* and Development of Metamodels for *STP*

The second order regression models are assumed to be;

$$Y_1 = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_{11}X_1^2 + \beta_{22}X_2^2 + \beta_{33}X_3^2 + \beta_{44}X_4^2 + \beta_{12}X_1X_2 + \beta_{13}X_1X_3 + \beta_{14}X_1X_4 + \beta_{23}X_2X_3 + \beta_{24}X_2X_4 + \beta_{34}X_3X_4 + \varepsilon$$

$$Y_2 = \delta_0 + \delta_1X_1 + \delta_2X_2 + \delta_3X_3 + \delta_4X_4 + \delta_{11}X_1^2 + \delta_{22}X_2^2 + \delta_{33}X_3^2 + \delta_{44}X_4^2 + \delta_{12}X_1X_2 + \delta_{13}X_1X_3 + \delta_{14}X_1X_4 + \delta_{23}X_2X_3 + \delta_{24}X_2X_4 + \delta_{34}X_3X_4 + \varepsilon \quad (5.17)$$

where β_0 and δ_0 are constant, and β_1, β_2, β_3, β_4, δ_1, δ_2, δ_3 and δ_4 are coefficients corresponding to main effects, and β_{12}, β_{13}, β_{14}, β_{23}, β_{24}, β_{34}, δ_{12}, δ_{13}, δ_{14}, δ_{23}, δ_{24} and δ_{34} are coefficients corresponding to two-factor interaction effects, β_{11}, β_{22}, β_{33}, β_{44}, δ_{11}, δ_{22}, δ_{33} and δ_{44} are coefficients corresponding to square effects, and ε is statistical error that have a normal distribution with mean zero and variance σ^2.

For fitting second order models a *CCD* is built by augmenting the two-level full factorial design with central runs and axial runs ($\alpha = 1$, means design is face-centered). In Table 5.40, the *CCD* used for fitting second order models for two responses is given. The values for two-level full factorial design points and axial points are average of 10 replications, and the values related to the central points are obtained by one replication.

After verifying that the residuals of the second order regression models fitted for Y_1 and Y_2 are normally, identically and independently distributed with zero mean and constant variance, the ANOVA tables are given in Table 5.41 and Table 5.42. Also estimated effects and coefficients (*EEC*) for two responses are denoted in Table 5.43 and 5.44 respectively.

Table 5.40 Simulation results for *STP* (*CCD*)

	X_1	X_2	X_3	X_4	Y_1	Y_2
1	-1	-1	-1	-1	523	35,3
2	1	-1	-1	-1	555	41,4
3	-1	1	-1	-1	621	46,1
4	1	1	-1	-1	653	57,2
5	-1	-1	1	-1	540	38,7
6	1	-1	1	-1	573	46,0
7	-1	1	1	-1	634	52,0
8	1	1	1	-1	665	66,4
9	-1	-1	-1	1	526	36,0
10	1	-1	-1	1	561	42,3
11	-1	1	-1	1	625	47,2
12	1	1	-1	1	657	58,9
13	-1	-1	1	1	545	39,6
14	1	-1	1	1	579	47,2
15	-1	1	1	1	640	53,6
16	1	1	1	1	671	68,9

	X_1	X_2	X_3	X_4	Y_1	Y_2
17	-1	0	0	0	587	45,6
18	1	0	0	0	621	56,3
19	0	-1	0	0	553	42,5
20	0	1	0	0	648	58,9
21	0	0	-1	0	598	48,1
22	0	0	1	0	614	54,6
23	0	0	0	-1	604	51,4
24	0	0	0	1	610	52,9
25	0	0	0	0	605	52,3
26	0	0	0	0	609	51,8
27	0	0	0	0	606	52,4
28	0	0	0	0	606	52,5
29	0	0	0	0	609	52,6
30	0	0	0	0	609	52,4
31	0	0	0	0	606	52,6

Table 5.41 ANOVA table for Y_1 response for *sorm* for *STP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	14	47592,8	47592,8	3399,5	2000,00	0,000
Linear	4	47033,2	47033,2	11758,3	8000,00	0,000
Square	4	531,9	531,9	133,0	90,34	0,000
Interaction	6	27,7	27,7	4,6	3,14	0,031
Residual Error	16	23,6	23,6	1,5		
Lack-of-Fit	10	4,7	4,7	0,5	0,15	0,995
Pure Error	6	18,9	18,9	3,1		
Total	30	47616,4				
S = 1,213 R-Sq = 99,9% R-Sq(adj) = 99,9%						

Table 5.42 ANOVA table for Y_2 response for *sorm* for *STP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	14	1868,7	1868,7	133,5	822,39	0,000
Linear	4	1720,2	1720,2	430,0	3000,00	0,000
Square	4	88,2	88,2	22,0	135,79	0,000
Interaction	6	60,4	60,4	10,1	61,99	0,000
Residual Error	16	2,6	2,6	0,2		
Lack-of-Fit	10	2,1	2,1	0,2	2,83	0,108
Pure Error	6	0,5	0,5	0,1		
Total	30	1871,3				
S = 0,4029 R-Sq = 99,9% R-Sq(adj) = 99,7%						

In Table 5.41 and Table 5.42, we see that linear, quadratic and two-way interaction effects are statistically significant ($P<0.05$), and the second order regression models are statistically significant ($P>0.05$) for both responses.

Table 5.43 *EEC* table for Y_1 response for *sorm* for *STP*

Term	Coef	SE Coef	T	P
Constant	606,896	0,360	1686,2620	0,000
X1	16,333	0,286	57,1170	0,000
X2	47,722	0,286	166,8820	0,000
X3	7,889	0,286	27,5870	0,000
X4	2,556	0,286	8,9370	0,000
*X1*X1*	-2,608	0,753	-3,4630	0,003
*X2*X2*	-6,108	0,753	-8,1100	0,000
*X3*X3*	-0,608	0,753	-0,8070	0,431
*X4*X4*	0,392	0,753	0,5200	0,610
*X1*X2*	-0,500	0,303	-1,6480	0,119
*X1*X3*	-0,125	0,303	-0,4120	0,686
*X1*X4*	0,250	0,303	0,8240	0,422
*X2*X3*	-1,125	0,303	-3,7090	0,002
*X2*X4*	0,000	0,303	0,0000	1,000
*X3*X4*	0,375	0,303	1,2360	0,234

Table 5.44 *EEC* table for Y_2 response for *sorm* for *STP*

Term	Coef	SE Coef	T	P
Constant	52,306	0,120	437,6630	0,000
X1	5,028	0,095	52,9470	0,000
X2	7,789	0,095	82,0250	0,000
X3	3,028	0,095	31,8860	0,000
X4	0,672	0,095	7,0790	0,000
*X1*X1*	-1,279	0,250	-5,1130	0,000
*X2*X2*	-1,529	0,250	-6,1130	0,000
*X3*X3*	-0,879	0,250	-3,5140	0,003
*X4*X4*	-0,079	0,250	-0,3150	0,757
*X1*X2*	1,575	0,101	15,6380	0,000
*X1*X3*	0,588	0,101	5,8330	0,000
*X1*X4*	0,125	0,101	1,2410	0,232
*X2*X3*	0,938	0,101	9,3080	0,000
*X2*X4*	0,200	0,101	1,9860	0,064
*X3*X4*	0,113	0,101	1,1170	0,280

The second order regression models are respectively;

$$Y_1 = 606{,}896 + 16{,}333\, x_1 + 47{,}722\, x_2 + 7{,}889\, x_3 + 2{,}556\, x_4 - 2{,}608\, x_1^2 - 6{,}108\, x_2^2 - 0{,}608\, x_3^2 + 0{,}392\, x_4^2 - 0{,}500\, x_1x_2 - 0{,}125\, x_1x_3 + 0{,}250\, x_1x_4 - 1{,}125\, x_2x_3 + 0{,}375\, x_3x_4$$

$Y_2 = 52,306 + 5,028\, x_1 + 7,789\, x_2 + 3,028\, x_3 + 0,672\, x_4 - 1,279\, x_1^2 - 1,529\, x_2^2 - 0,879\, x_3^2 - 0,079\, x_4^2 + 1,575\, x_1x_2 + 0,588\, x_1x_3 + 0,125\, x_1x_4 + 0,938\, x_2x_3 + 0,200\, x_2x_4 + 0,113\, x_3x_4$ (5.18)

After building regression metamodels for both responses, they are verified and validated.

5.3.2.1.2 Verification and Validation of *STP* Metamodels

For verification, lack-of-fit tests are applied for both metamodels. As can be seen from Table 5.41 and Table 5.42, the metamodels have no statistically significant lack-of-fit (P>0.05) for both responses with 5% significance level ($\alpha = 0.05$). R^2 is 99.9%, and R^2-*adjusted* is 99.9% for response Y_1, and R^2 is 99.9%, and R^2-*adjusted* is 99.7% for response Y_2.

Table 5.45 Metamodels validation for *STP*

	Input factors				Simulation model		Metamodel		ARE (%)	
	X_1	X_2	X_3	X_4	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
1	-0,6	-0,2	-0,2	-0,6	578	46,5	583,4	46,52	0,93	0,04
2	0,2	-0,2	-0,2	-0,6	590	50,5	597,3	50,54	1,23	0,08
3	-0,2	0,2	-0,6	-0,2	602	50,4	607,6	50,38	0,93	0,04
4	0,2	0,6	-0,6	-0,2	627	54,9	631,4	54,88	0,71	0,04
5	-0,2	-0,2	0,6	-0,6	592	50,5	596,9	50,58	0,82	0,15
6	0,6	-0,6	0,2	-0,2	583	49,2	586,2	49,47	0,56	0,54
7	-0,2	0,6	0,6	-0,6	626	56,4	632,6	56,42	1,06	0,04
8	0,2	0,2	0,2	-0,6	626	55,3	619,4	54,97	1,06	0,60
9	-0,6	-0,2	-0,2	0,6	580	47,1	586,2	47,16	1,07	0,12
10	0,6	-0,6	-0,6	0,6	575	47,3	581,4	47,37	1,12	0,15
11	-0,2	0,6	-0,2	0,2	623	54,7	629,0	54,60	0,97	0,18
12	0,6	0,6	-0,6	0,6	635	57,4	639,1	57,32	0,65	0,14
13	-0,6	-0,6	0,6	0,2	565	45,2	570,7	45,23	1,00	0,06
14	0,6	-0,2	0,2	0,2	602	53,8	608,2	53,80	1,03	0,00
15	-0,2	0,6	0,6	0,6	629	57,5	635,9	57,42	1,10	0,13
16	0,2	0,2	0,6	0,2	629	56,6	624,3	56,66	0,75	0,11

For validation, 16 (2^4) points (because there are 4 input factors with 2 level and we want to see if metamodel is valid in the entire experimentation region) that selected randomly in experimentation region, and from different design points, are used. These 16 points are shown in Table 5.45. The simulation model is run at these randomly selected points with 10 replications for each point. The values that are denoted in the Simulation Model column of Table 5.45 are the average of these 10 replications. Also the values denoted in Metamodel column are predicted by fitted second order regression metamodels. Since the entire *ARE* values are smaller than 3% the second order metamodels can be used for prediction.

After verification and validation of regression metamodels, Derringer-Suich multi-response optimization procedure is used for optimization.

5.3.2.1.3 Derringer-Suich Multi-Response Optimization Procedure for *STP*

In the current problem, since the target value for the fullness rate of the carriages response (Y_2) is 50%, the lower and upper limits are determined as 45% and 55%, respectively. To decide the upper and lower limits on the average passenger time spent in the metro-line response (Y_1) the simulation model is run 20 times at the low levels of all input factors. The lower limit is defined as the average of these 20 values, and is found as 522 seconds. To determine the upper limit, simulation model is run 20 times at the high levels of all input factors, and the average value, 671 seconds, is found as upper limit.

We will use three values 0.1, 1 and 10 for t and s. Therefore, nine situations (combination of weights) appear to be evaluated.

For each combination of weights, the lower, target, upper values and weights for two responses, and the optimum conditions with coded variables, and also corresponding predicted responses at these coded values which are calculated from the fitted models, in addition the individual desirability values and the composite desirability values at optimum conditions are shown in Table 5.46.

Table 5.46 Derringer-Suich optimization method results for *STP*

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	522	522	671	0,1	1
Y2	Target	45	50	55	0,1	1

Global Solution

X1	=	0,99568
X2	=	-0,82360
X3	=	0,99671
X4	=	0,99825

Predicted Responses

Y1	=	589,158;	desirability	=	0,94184
Y2	=	49,923;	desirability	=	0,99845

Composite Desirability = 0,96973

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	522	522	671	0,1	1
Y2	Target	45	50	55	1,0	1

Global Solution

X1	=	0,99790
X2	=	-0,81820
X3	=	0,99840
X4	=	0,99915

Predicted Responses

Y1	=	589,504;	desirability	=	0,94145
Y2	=	50,000;	desirability	=	0,99997

Composite Desirability = 0,97027

Parameters

	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	522	522	671	0,1	1
Y2	Target	45	50	55	10,0	1

Global Solution

X1	=	1,00000
X2	=	-0,75472
X3	=	1,00000
X4	=	-0,17914

Predicted Responses

Y1	=	588,951;	desirability	=	0,94208
Y2	=	50,007;	desirability	=	0,98574

Composite Desirability = 0,96367

Table 5.46 ***(Continued)***

```
Parameters
                    Goal      Lower     Target     Upper    Weight    Import
Y1               Minimum        522        522       671       1,0         1
Y2                Target         45         50        55       0,1         1
Global Solution
X1          =  0,63587
X2          = -1,00000
X3          =  0,65441
X4          =  0,90807
Predicted Responses
Y1          =  571,313; desirability =  0,66904
Y2          =   46,405; desirability =  0,88080
Composite Desirability =  0,76766
```

```
Parameters
                    Goal      Lower     Target     Upper    Weight    Import
Y1               Minimum        522        522       671         1         1
Y2                Target         45         50        55         1         1
Global Solution
X1          =  1,00000
X2          = -0,75656
X3          =  1,00000
X4          = -0,13414
Predicted Responses
Y1          =  588,986; desirability =  0,55043
Y2          =   50,019; desirability =  0,99615
Composite Desirability =  0,74048
```

```
Parameters
                    Goal      Lower     Target     Upper    Weight    Import
Y1               Minimum        522        522       671         1         1
Y2                Target         45         50        55        10         1
Global Solution
X1          =  0,99790
X2          = -0,81820
X3          =  0,99840
X4          =  0,99915
Predicted Responses
Y1          =  589,504; desirability =  0,54696
Y2          =   50,000; desirability =  0,99972
Composite Desirability =  0,73946
```

Table 5.46 ***(Continued)***

Parameters						
	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	522	522	671	10,0	1
Y2	Target	45	50	55	0,1	1
Global Solution						
X1	= 0,39478					
X2	= -1,00000					
X3	= 1,00000					
X4	= -0,11385					
Predicted Responses						
Y1	= 567,322;	desirability =	0,02661			
Y2	= 45,522;	desirability =	0,79774			
Composite Desirability =	0,14569					
Parameters						
	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	522	522	671	10	1
Y2	Target	45	50	55	1	1
Global Solution						
X1	= 1,00000					
X2	= -1,00000					
X3	= 0,65333					
X4	= 0,66578					
Predicted Responses						
Y1	= 575,044;	desirability =	0,01227			
Y2	= 46,948;	desirability =	0,38957			
Composite Desirability =	0,06914					
Parameters						
	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	522	522	671	10	1
Y2	Target	45	50	55	10	1
Global Solution						
X1	= 1,00000					
X2	= -0,75457					
X3	= 1,00000					
X4	= -0,18159					
Predicted Responses						
Y1	= 588,951;	desirability =	0,00256			
Y2	= 50,007;	desirability =	0,98584			
Composite Desirability =	0,05027					

Since the optimum levels of input factors are found by using the metamodels, confirmatory simulation runs are needed at optimum input levels. The simulation model responses shown in Table 5.47 are the average values of 10 replications.

In this table, first nine rows reflect the global optimum solutions found by Derringer-Suich multi-response optimization method according to weight combinations, and the last row belongs to current situation.

Table 5.47 Results of the confirmatory runs for *STP*

	Weight		Composite	Natural variables				Metamodel		Simulation	
	Y_1	Y_2	desirability	X_1	X_2	X_3	X_4	Y_1	Y_2	Y_1	Y_2
1	0,1	0,1	0,96973	599	326	600	900	589,2	49,92	592	49,6
2	0,1	1	0,97027	600	327	600	900	589,5	50,00	590	49,9
3	0,1	10	0,96367	600	337	600	723	589,0	50,01	593	50,0
4	1	0,1	0,76766	545	300	548	886	571,3	46,41	571	45,8
5	1	1	0,74048	600	337	600	730	589,0	50,02	594	50,2
6	1	10	0,73946	600	327	600	900	589,5	50,00	592	49,9
7	10	0,1	0,14569	509	300	600	733	567,3	45,52	571	45,2
8	10	1	0,06914	600	300	548	850	575,0	46,95	575	46,7
9	10	10	0,05027	600	337	600	723	589,0	50,01	593	50,1
10	Current		X	600	450	600	900	X	X	632	59,8

Table 5.48 Factor values of confirmatory runs for *STP*

	Coded values				Natural values (second)				Natural values (minutes)			
	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4
1	1,00	-0,82	1,00	1,00	599	326	600	900	10,0	5,4	10,0	15,0
2	1,00	-0,82	1,00	1,00	600	327	600	900	10,0	5,5	10,0	15,0
3	1,00	-0,75	1,00	-0,18	600	337	600	723	10,0	5,6	10,0	12,1
4	0,64	-1,00	0,65	0,91	545	300	548	886	9,1	5,0	9,1	14,8
5	1,00	-0,76	1,00	-0,13	600	337	600	730	10,0	5,6	10,0	12,2
6	1,00	-0,82	1,00	1,00	600	327	600	900	10,0	5,5	10,0	15,0
7	0,39	-1,00	1,00	-0,11	509	300	600	733	8,5	5,0	10,0	12,2
8	1,00	-1,00	0,65	0,67	600	300	548	850	10,0	5,0	9,1	14,2
9	1,00	-0,75	1,00	-0,18	600	337	600	723	10,0	5,6	10,0	12,0
10	1,00	0,00	1,00	1,00	600	450	600	900	10,0	7,5	10,0	15,0

In Table 5.48, factor values of confirmatory runs for *STP* are given in the first nine rows and the last row belongs to current situation.

First, the fullness rate of the carriages response (Y_2) is checked. In Table 5.47, we see that the fullness rate requirement 50% is provided by the first, second, third, fifth, sixth and ninth rows. After determining the natural factor levels providing the fullness rate requirement, the minimum value for the average passenger time spent in the metro-line (Y_1) is searched, and it is seen that the minimum Y_1 (590 seconds) is obtained with the natural factor levels in the 2nd row. As a result, one optimum point exist. For $X_1 = 10.0, X_2 = 5.5, X_3 = 10.0, X_4 = 15.0$ factor levels the obtained values for the average passenger time spent in the metro-line (Y_1), and for the fullness rate of carriages (Y_2) responses are respectively 590 seconds and 49.9 percent. This optimum point's factor levels are demonstrated in grey colour in Tables 5.47 and 5.48.

As can be seen from Table 5.47 after optimization study. the fullness rate of carriages (Y_2) is decreased by 9.9 percent and the average passenger time spent in the metro-line (Y_1) is decreased by 42 seconds (6.65 percentage) respect to current values.

5.4 Sunday Problem (*SNP*)

5.4.1 Estimation Process

5.4.1.1 Phase Zero

The objective of *SNP* is to find the levels of 3 input factors that related to the 3 time periods from 06:00 a.m. to 00:00 a.m. and which minimize the average passenger time spent in the metro-line with the requirement as the fifty percent fullness rate of the carriages. The built simulation model is explained in chapter four.

Generally at phase zero, a screening experiment is made for investigating potential input factors, which are thought to be important in the response surface study, and for determining important factors. Because the input factors are 3 headways and none of them can be eliminated we skip factor screening processes.

The input factors are respectively;

X_1: The headway for the first time period from 06:00 a.m. to 09:00 a.m.

X_2: The headway for the second time period from 09:00 a.m. to 20:00 p.m.

X_3: The headway for the third time period from 20:00 p.m. to 00:00 a.m.

The output responses are respectively;

Y_1: The average passenger time that is spent in the metro-line (in second)

Y_2: The fullness rate of the carriages (in percentage)

5.4.1.2 Phase One

Low and high level of input factors for *SNP*, time period durations, and the natural values are given in Table 5.49.

Table 5.49 Low and high level of input factors for *SNP*

Input factor	Time Period	Duration		Low level		High level	
		(hour)	(second)	Coded	Natural (second)	Coded	Natural (second)
X_1	06:00 - 09:00	3,0	10800	-1	600	1	900
X_2	09:00 - 20:00	11,0	39600	-1	600	1	900
X_3	20:00 - 00:00	4,0	14400	-1	600	1	900

The first order regression models with two-factor interactions are assumed to be;

$$Y_1 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 + \varepsilon$$

$$Y_2 = \delta_0 + \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3 + \delta_{12} X_1 X_2 + \delta_{13} X_1 X_3 + \delta_{23} X_2 X_3 + \varepsilon \qquad (5.19)$$

where β_0 and δ_0 are constant, and β_1, β_2, β_3, δ_1, δ_2 and δ_3 are coefficients corresponding to main effects, and β_{12}, β_{13}, β_{23}, δ_{12}, δ_{13} and δ_{23} are coefficients corresponding to two-factor interaction effects and ε is statistical error that have a normal distribution with mean zero and variance σ^2.

5.4.1.2.1 Two-level Full Factorial Design

For fitting a first order regression model a two-level full factorial design (2^3) with central runs is designed and then the simulation model is run 10 times at each design point, and response values for each design points are found.

Although the values for two-level full factorial design points, which are denoted in Table 5.50, are average of 10 replications, the values related to the central points are obtained by one replication.

Table 5.50 Simulation results for *SNP* (2^3 design with 5 central runs)

	X_1	X_2	X_3	Y_1	Y_2
1	-1	-1	-1	665	34,2
2	1	-1	-1	675	36,3
3	-1	1	-1	788	43,0
4	1	1	-1	799	46,2
5	-1	-1	1	681	36,8
6	1	-1	1	693	39,2
7	-1	1	1	803	47,2

	X_1	X_2	X_3	Y_1	Y_2
8	1	1	1	813	51,1
9	0	0	0	738	42,4
10	0	0	0	739	42,6
11	0	0	0	735	42,9
12	0	0	0	735	42,3
13	0	0	0	736	43,0

After verifying that the residuals are normally, identically and independently distributed with zero mean and constant variance for the Y_1 and Y_2 responses, Variance Analysis is used to see if the main effects, two-factor interactions, curvature and lack-of-fit are statistically significant. Table 5.51 is the Analysis of Variance Table (ANOVA) for the first response Y_1, and Table 5.52 for the second response Y_2. In addition, estimated effects and coefficients (*EEC*) for two responses are given in Table 5.53 and 5.54, respectively.

Table 5.51 ANOVA table for Y_1 response for *form* for *SNP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	30617,4	30617,4	10205,8	1000,00	0,000
2-Way Interactions	3	3,4	3,4	1,1	0,16	0,920
Residual Error	6	42,5	42,5	7,1		
Curvature	1	28,2	28,2	28,2	9,83	0,026
Lack of Fit	1	1,1	1,1	1,1	0,34	0,591
Pure Error	4	13,2	13,2	3,3		
Total	12	30663,2				
S = 2,257 R-Sq = 99,9% R-Sq(adj) = 99,8%						

In Table 5.51, since the *P*-value for the lack-of-fit test is greater than the significance level ($\alpha = 0.05$), the lack-of-fit is not statistically significant, that is, the first order regression model for Y_1 response for *WAP* is adequate. The main effects and curvature are statically significant (*P*-value<0.05), two-way interactions are not statistically significant (*P*-value>0.05).

Table 5.52 ANOVA table for Y_2 response for *form* for *SNP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	253,6	253,6	84,5	179,26	0,000
2-Way Interactions	3	2,6	2,6	0,9	1,83	0,242
Residual Error	6	2,8	2,8	0,5		
Curvature	1	2,4	2,4	2,4	31,09	0,003
Lack of Fit	1	0,0	0,0	0,0	0,22	0,667
Pure Error	4	0,4	0,4	0,1		
Total	12	259,0				
S = 0,7760 R-Sq = 97,9% R-Sq(adj) = 97,2%						

Table 5.52 shows that the lack-of-fit and two-way interactions are not statically significant (P>0.05), curvature and main effects are statistically significant (P<0.05).

Table 5.53 *EEC* table for Y_1 response for *form* for *SNP*

Term	Effect	Coef	SE Coef	T	P
Constant		738,462	0,7380	1000,64	0,000
X1	10,750	5,375	0,9408	5,71	0,001
X2	122,250	61,125	0,9408	64,97	0,000
X3	15,750	7,875	0,9408	8,37	0,000
*X1*X2*	-0,250	-0,125	0,9408	-0,13	0,899
*X1*X3*	0,250	0,125	0,9408	0,13	0,899
*X2*X3*	-1,250	-0,625	0,9408	-0,66	0,531

Table 5.54 *EEC* table for Y_2 response for *form* for *SNP*

Term	Effect	Coef	SE Coef	T	P
Constant		42,092	0,1905	221,01	0,000
X1	2,900	1,450	0,2428	5,97	0,001
X2	10,250	5,125	0,2428	21,11	0,000
X3	3,650	1,825	0,2428	7,52	0,000
*X1*X2*	0,650	0,325	0,2428	1,34	0,229
*X1*X3*	0,250	0,125	0,2428	0,51	0,625
*X2*X3*	0,900	0,450	0,2428	1,85	0,113

As a result, the fitted first order models for both responses are respectively;

$$Y_1 = 738{,}462 + 5{,}375\, x_1 + 61{,}125\, x_2 + 7{,}875\, x_3 - 0{,}125\, x_1x_2 + 0{,}125\, x_1x_3 - 0{,}625\, x_2x_3$$

$$Y_2 = 42{,}092 + 1{,}450\, x_1 + 5{,}125\, x_2 + 1{,}825\, x_3 + 0{,}325 x_1x_2 + 0{,}125\, x_1x_3 + 0{,}450\, x_2x_3 \qquad (5.20)$$

Because the true response surface usually exhibits curvature near the optimum, it is understood that the determined region of experimentation is near the region of optimum. As mentioned in chapter two, the second of *RSM* begins. Therefore, second order models can be built.

5.4.2 Optimization Process

5.4.2.1 Phase Two

5.4.2.1.1 *CCD* and Development of Metamodels for *SNP*

The second order regression models are assumed to be;

$$Y_1 = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_{11}X_1^2 + \beta_{22}X_2^2 + \beta_{33}X_3^2 + \beta_{12}X_1X_2 + \beta_{13}X_1X_3 + \beta_{23}X_2X_3 + \varepsilon$$

$$Y_2 = \delta_0 + \delta_1X_1 + \delta_2X_2 + \delta_3X_3 + \delta_{11}X_1^2 + \delta_{22}X_2^2 + \delta_{33}X_3^2 + \delta_{12}X_1X_2 + \delta_{13}X_1X_3 + \delta_{23}X_2X_3 + \varepsilon \qquad (5.21)$$

where β_0 and δ_0 are constant, and β_1, β_2, β_3, δ_1, δ_2 and δ_3 are coefficients corresponding to main effects, and β_{12}, β_{13}, β_{23}, δ_{12}, δ_{13}, and δ_{23} are coefficients corresponding to two-factor interaction effects, β_{11}, β_{22}, β_{33}, δ_{11}, δ_{22} and δ_{33} are coefficients corresponding to square effects, and ε is statistical error that have a normal distribution with mean zero and variance σ^2.

For fitting second order models a *CCD* is built by augmenting the two-level full factorial design with central runs and axial runs ($\alpha = 1$, means design is face-centered). In Table 5.55, the *CCD* used for fitting second order models for two responses is given. The values for two-level full factorial design points and axial points are average of 10 replications, and the values related to the central points are obtained by one replication.

Table 5.55 Simulation results for *SNP* (*CCD*)

	X_1	X_2	X_3	Y_1	Y_2
1	-1	-1	-1	665	34,2
2	1	-1	-1	675	36,3
3	-1	1	-1	788	43,0
4	1	1	-1	799	46,2
5	-1	-1	1	681	36,8
6	1	-1	1	693	39,2
7	-1	1	1	803	47,2
8	1	1	1	813	51,1
9	-1	0	0	732	41,1
10	1	0	0	745	43,8

	X_1	X_2	X_3	Y_1	Y_2
11	0	-1	0	676	37,1
12	0	1	0	801	47,5
13	0	0	-1	730	40,5
14	0	0	1	748	44,3
15	0	0	0	735	42,7
16	0	0	0	738	42,4
17	0	0	0	739	42,6
18	0	0	0	735	42,9
19	0	0	0	735	42,3
20	0	0	0	736	43,0

After verifying that the residuals of the second order regression models fitted for Y_1 and Y_2 are normally, identically and independently distributed with zero mean and constant variance, the ANOVA tables are given in Table 5.56 and Table 5.57. Also estimated effects and coefficients (*EEC*) for two responses are denoted in Table 5.58 and 5.59 respectively.

Table 5.56 ANOVA table for Y_1 response for *sorm* for *SNP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	38705,0	38705,0	4300,6	1000,00	0,000
Linear	3	38669,3	38669,3	12889,8	4000,00	0,000
Square	3	32,3	32,3	10,8	3,65	0,052
Interaction	3	3,4	3,4	1,1	0,38	0,769
Residual Error	10	29,5	29,5	3,0		
Lack-of-Fit	5	14,2	14,2	2,8	0,93	0,533
Pure Error	5	15,3	15,3	3,1		
Total	19	38734,5				
S = 1,719 R-Sq = 99,9% R-Sq(adj) = 99,9%						

In Table 5.56, we see that linear effect is statistically significant (P<0.05), two-way interaction and quadratic effects are not statistically significant (P>0.05), also the second order regression model is statistically significant (P>0.05).

Table 5.57 ANOVA table for Y_2 response for *sorm* for *SNP*

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	324,1	324,1	36,0	831,08	0,000
Linear	3	318,5	318,5	106,2	2000,00	0,000
Square	3	3,1	3,1	1,0	23,49	0,000
Interaction	3	2,6	2,6	0,9	19,92	0,000
Residual Error	10	0,4	0,4	0,0		
Lack-of-Fit	5	0,1	0,1	0,0	0,16	0,969
Pure Error	5	0,4	0,4	0,1		
Total	19	324,6				
S = 0,2082 R-Sq = 99,9% R-Sq(adj) = 99,7%						

In Table 5.57, we see that linear, quadratic and two-way interaction effects are statistically significant ($P<0.05$), and the second order regression model is statistically significant ($P>0.05$).

Table 5.58 *EEC* table for Y_1 response for *sorm* for *SNP*

Term	Coef	SE Coef	T	P
Constant	736,873	0,591	1247,2550	0,000
X1	5,600	0,544	10,3040	0,000
X2	61,400	0,544	112,9810	0,000
X3	8,100	0,544	14,9050	0,000
*X1*X1*	0,818	1,036	0,7900	0,448
*X2*X2*	0,818	1,036	0,7900	0,448
*X3*X3*	1,318	1,036	1,2720	0,232
*X1*X2*	-0,125	0,608	-0,2060	0,841
*X1*X3*	0,125	0,608	0,2060	0,841
*X2*X3*	-0,625	0,608	-1,0290	0,328

Table 5.59 *EEC* table for Y_2 response for *sorm* for *SNP*

Term	Coef	SE Coef	T	P
Constant	42,665	0,072	596,1640	0,000
X1	1,430	0,066	21,7220	0,000
X2	5,140	0,066	78,0790	0,000
X3	1,840	0,066	27,9510	0,000
*X1*X1*	-0,236	0,126	-1,8830	0,089
*X2*X2*	-0,386	0,126	-3,0780	0,012
*X3*X3*	-0,286	0,126	-2,2810	0,046
*X1*X2*	0,325	0,074	4,4160	0,001
*X1*X3*	0,125	0,074	1,6980	0,120
*X2*X3*	0,450	0,074	6,1140	0,000

As a result, the second order regression models are respectively;

$$Y_1 = 736{,}873 + 5{,}600\,x_1 + 61{,}400\,x_2 + 8{,}100\,x_3 + 0{,}818\,x_1^2 + 0{,}818\,x_2^2 + 1{,}318\,x_3^2 - 0{,}125\,x_1x_2 + 0{,}125\,x_1x_3 - 0{,}625\,x_2x_3$$

$$Y_2 = 42{,}665 + 1{,}430\,x_1 + 5{,}140\,x_2 + 1{,}840\,x_3 - 0{,}236\,x_1^2 - 0{,}386\,x_2^2 - 0{,}286\,x_3^2 + 0{,}325\,x_1x_2 + 0{,}125\,x_1x_3 + 0{,}450\,x_2x_3 \quad (5.22)$$

After building regression metamodels for both responses, they are verified and validated.

5.4.2.1.2 Verification and Validation of *SNP* Metamodels

For verification, lack-of-fit tests are applied for both metamodels. As can be seen from Table 5.56 and Table 5.57, the metamodels have no statistically significant lack-of-fit (P>0.05) for both responses with 5% significance level ($\alpha = 0.05$). R^2 is 99.9%, and R^2-*adjusted* is 99.9% for response Y_1, and R^2 is 99.9%, and R^2-*adjusted* is 99.7% for response Y_2.

For validation, 8 (2^3) points (because there are 3 input factors with 2 level and we want to see if metamodel is valid in the entire experimentation region) that selected randomly in experimentation region, and from different design points, are used. These 8 points are shown in Table 5.60. The simulation model is run at these randomly selected points with 10 replications for each point. The values that are denoted in the Simulation Model column of Table 5.60 are the average of these 10 replications. Also the values denoted in Metamodel column are predicted by fitted second order regression metamodels. Since the entire *ARE* values are smaller than 3% the second order metamodels can be used for prediction.

Table 5.60 Metamodels validation for *SNP*

	Input factors			Simulation model		Metamodel		ARE (%)	
	X_1	X_2	X_3	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
1	-0,2	-0,2	-0,6	737	40,1	719,1	40,20	2,43	0,25
2	0,2	-0,2	-0,2	725	41,4	724,2	41,52	0,11	0,29
3	-0,2	0,6	-0,2	774	44,9	771,4	44,85	0,33	0,12
4	0,2	0,6	-0,2	774	45,6	773,6	45,49	0,05	0,25
5	-0,2	-0,2	0,6	734	42,2	728,9	42,27	0,69	0,17
6	0,6	-0,2	0,2	736	42,6	730,0	42,71	0,81	0,25
7	-0,2	0,6	0,2	792	45,8	774,5	45,68	2,21	0,26
8	0,6	0,2	0,6	758	45,7	758,1	45,59	0,02	0,24

After verification and validation of regression metamodels, Derringer-Suich multi-response optimization procedure is used for optimization.

5.4.2.1.3 Derringer-Suich Multi-Response Optimization Procedure for *SNP*

In the current problem, since the target value for the fullness rate of the carriages response (Y_2) is 50%, the lower and upper limits are determined as 45% and 55%, respectively. To decide the upper and lower limits on the average passenger time spent in the metro-line response (Y_1) the simulation model is run 20 times at the low levels of all input factors. The lower limit is defined as the average of these 20 values, and is found as 663 seconds. To determine the upper limit, simulation model is run 20 times at the high levels of all input factors, and the average value, 813 seconds, is found as upper limit.

We will use three values 0.1, 1 and 10 for t and s. Therefore, nine situations (combination of weights) appear to be evaluated.

For each combination of weights, the lower, target, upper values and weights for two responses, and the optimum conditions with coded variables, and also corresponding predicted responses at these coded values which are calculated from the fitted models, in addition the individual desirability values and the composite desirability values at optimum conditions are shown in Table 5.61.

Table 5.61 Derringer-Suich optimization method results for *SNP*

```
Parameters
                 Goal      Lower    Target    Upper    Weight    Import
Y1            Minimum        663       663      813       0,1         1
Y2             Target         45        50       55       0,1         1
Global Solution
X1            =  1,00000
X2            =  0,43763
X3            =  1,00000
Predicted Responses
Y1            =  779,533; desirability =  0,86070
Y2            =   48,051; desirability =  0,95182
Composite Desirability =  0,90511
```

```
Parameters
                 Goal      Lower    Target    Upper    Weight    Import
Y1            Minimum        663       663      813       0,1         1
Y2             Target         45        50       55       1,0         1
Global Solution
X1            =  1,00000
X2            =  0,79593
X3            =  1,00000
Predicted Responses
Y1            =  801,626; desirability =  0,77265
Y2            =   50,000; desirability =  1,00000
Composite Desirability =  0,87900
```

```
Parameters
                 Goal      Lower    Target    Upper    Weight    Import
Y1            Minimum        663       663      813       0,1         1
Y2             Target         45        50       55      10,0         1
Global Solution
X1            =  1,00000
X2            =  0,79593
X3            =  1,00000
Predicted Responses
Y1            =  801,626; desirability =  0,77265
Y2            =   50,000; desirability =  1,00000
Composite Desirability =  0,87901
```

Table 5.61 ***(Continued)***

Parameters						
	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	663	663	813	1,0	1
Y2	Target	45	50	55	0,1	1

Global Solution

X1 = 1,00000

X2 = -0,00147

X3 = 1,00000

Predicted Responses

Y1 = 752,745; desirability = 0,40170

Y2 = 45,528; desirability = 0,79869

Composite Desirability = 0,56642

Parameters						
	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	663	663	813	1	1
Y2	Target	45	50	55	1	1

Global Solution

X1 = 1,00000

X2 = 0,43992

X3 = 1,00000

Predicted Responses

Y1 = 779,674; desirability = 0,22218

Y2 = 48,064; desirability = 0,61284

Composite Desirability = 0,36899

Parameters						
	Goal	Lower	Target	Upper	Weight	Import
Y1	Minimum	663	663	813	1	1
Y2	Target	45	50	55	10	1

Global Solution

X1 = 1,00000

X2 = 0,79593

X3 = 1,00000

Predicted Responses

Y1 = 801,626; desirability = 0,07583

Y2 = 50,000; desirability = 1,00000

Composite Desirability = 0,27537

Table 5.61 ***(Continued)***

```
Parameters
                    Goal      Lower     Target      Upper     Weight     Import
Y1               Minimum        663        663        813       10,0          1
Y2                Target         45         50         55        0,1          1
Global Solution
X1            =  1,00000
X2            =  0,00914
X3            =  1,00000
Predicted Responses
Y1            =  753,388; desirability =  0,00010
Y2            =   45,591; desirability =  0,80770
Composite Desirability =  0,00891
```

```
Parameters
                    Goal      Lower     Target      Upper     Weight     Import
Y1               Minimum        663        663        813         10          1
Y2                Target         45         50         55          1          1
Global Solution
X1            =  1,00000
X2            =  0,52796
X3            =  1,00000
Predicted Responses
Y1            =  785,083; desirability =  0,00000
Y2            =   48,552; desirability =  0,71040
Composite Desirability =  0,00019
```

```
Parameters
                    Goal      Lower     Target      Upper     Weight     Import
Y1               Minimum        663        663        813         10          1
Y2                Target         45         50         55         10          1
Global Solution
X1            =  1,00000
X2            =  0,39150
X3            =  1,00000

Predicted Responses
Y1            =  776,704; desirability =  0,00000
Y2            =   47,793; desirability =  0,00296
Composite Desirability =  0,00005
```

Since the optimum levels of input factors are found by using the metamodels, confirmatory simulation runs are needed at optimum input levels. The simulation model responses shown in Table 5.62 are the average values of 10 replications.

Table 5.62 Results of the confirmatory runs for *SNP*

	Weight		Composite	Natural variables			Metamodel		Simulation	
	Y_1	Y_2	desirability	X_1	X_2	X_3	Y_1	Y_2	Y_1	Y_2
1	0,1	0,1	0,90511	900	816	900	779,5	48,05	779	48,2
2	0,1	1	0,87900	900	869	900	801,6	50,00	799	49,9
3	0,1	10	0,87901	900	869	900	801,6	50,00	799	50,1
4	1	0,1	0,56642	900	750	900	752,7	45,53	757	45,7
5	1	1	0,36899	900	816	900	779,7	48,06	778	48,2
6	1	10	0,27537	900	869	900	801,6	50,00	800	50,0
7	10	0,1	0,00891	900	751	900	753,4	45,59	757	45,6
8	10	1	0,00019	900	829	900	785,1	48,55	783	48,6
9	10	10	0,00005	900	809	900	776,7	47,79	776	47,8
10	Current		X	900	600	900	X	X	694	39,3

In this table, first nine rows reflect the global optimum solutions found by Derringer-Suich multi-response optimization method according to weight combinations, and the last row belongs to current situation.

Table 5.63 Factor values of confirmatory runs for *SNP*

	Coded values			Natural values (second)			Natural values (minutes)		
	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
1	1,00	0,44	1,00	900	816	900	15,0	13,6	15,0
2	1,00	0,80	1,00	900	869	900	15,0	14,5	15,0
3	1,00	0,80	1,00	900	869	900	15,0	14,5	15,0
4	1,00	0,00	1,00	900	750	900	15,0	12,5	15,0
5	1,00	0,44	1,00	900	816	900	15,0	13,6	15,0
6	1,00	0,80	1,00	900	869	900	15,0	14,5	15,0
7	1,00	0,01	1,00	900	751	900	15,0	12,5	15,0
8	1,00	0,53	1,00	900	829	900	15,0	13,8	15,0
9	1,00	0,39	1,00	900	809	900	15,0	13,5	15,0
10	1,00	-1,00	1,00	900	600	900	15,0	10,0	15,0

In Table 5.63, factor values of confirmatory runs for *SNP* are given in the first nine rows and the last row belongs to current situation.

First, the fullness rate of the carriages response (Y_2) is checked. In Table 5.62, we see that the fullness rate requirement 50% is provided by the second, third and sixth rows. After determining the natural factor levels that provides the fullness rate requirement, the minimum value for the average passenger time spent in the metro-line (Y_1) is searched, and it is seen that the natural factor levels of these three rows are same.

The small differences between responses are caused by random numbers, so it means 30 replications are done at these factor levels. The averages of these 30 values are used for finding optimum levels. For $X_1 = 15.0, X_2 = 14.5, X_3 = 15.0$ factor levels the obtained values for the average passenger time spent in the metro-line (Y_1), and for the fullness rate of carriages (Y_2) responses are respectively 799 seconds and 50.0 percent. This optimum point's factor levels are demonstrated in grey colour in Tables 5.62 and 5.63.

As can be seen from Table 5.62 after optimization study the fullness rate of carriages (Y_2) is increased by 10.7 percent and the average passenger time spent in the metro-line (Y_1) is increased by 105 seconds (15.13 percent). Although the current average passenger time spent in the metro-line (Y_1) value is less than optimum result, it violates the fullness rate of carriages requirement (50%).

CHAPTER SIX
CONCLUSION

In this thesis, for an underground public transport system in Izmir city, the aim is to find the optimum headways (input factors), to minimize the average passenger time spent in the metro-line (the first response) with the requirement of fifty percent (the second response) fullness rate of the carriages.

First, the problem in Izmir Metro is defined in detail. After explaining the working logics of the system and simulation model, input data for model, made assumptions before coding phase of the simulation model, flowcharts for occurred events, some important attributes and variables that are used in model are given. After that, flowcharts that denote modelling logic of the Halkapinar station are demonstrated. Then, the verification and validation techniques that are used for simulation model are given.

The built simulation model is generic model that easily changes to suit the changes in headways, in the length of time periods, and in the number of carriages. Simulation model can reply what-if questions as;

- What will be the values of responses if headways changes?
- What will be the values of responses if the carriage numbers of trains change? And also the carriage numbers of trains can be different in two consecutive time periods. Namely, the carriage numbers must be same for a certain time period, but can be different through a day.
- What will be the values of responses if the lengths and number of time periods change?

After building the simulation model, the optimization points are searched for four problems, which are weekday morning (*WMP*), weekday afternoon (*WAP*), Saturday

(*STP*) and Sunday (*SNP*) problems. First, metamodels are developed for both responses for four problems. After verifying and validating the metamodels Derringer-Suich multi-response optimization procedure is used for searching factor levels. The developed metamodels are flexible for searching the factor levels for different fullness rate requirements. Namely, the current fullness rate requirement is 50% for four problems, but it can be changed, also required values can be different for each problem.

As a result, two optimum points found for *WMP*. For $X_1 = 9.8, X_2 = 9.9, X_3 = 7.1, X_4 = 9.0, X_5 = 7.8$ factor levels and $X_1 = 10.0, X_2 = 10.0, X_3 = 7.5, X_4 = 7.8, X_5 = 7.7$ factor levels the obtained values for the average passenger time spent in the metro-line (Y_1), and for the fullness rate of carriages (Y_2) responses are respectively 625 seconds and 50,3 percent. After optimization study the fullness rate of carriages (Y_2) is increased by 4.1 % and the average passenger time spent in the metro-line (Y_1) is increased by 12 seconds (1.96 %) respect to current values. Although the current average passenger time spent in the metro-line (Y_1) value is less than optimum result, it violates the fullness rate of carriages requirement (50 %).

Then, an optimum point is found for *WAP*. For $X_1 = 7.3$, $X_2 = 6.2$, $X_3 = 6.8$, $X_4 = 10.0, X_5 = 10.5$ factor levels the obtained values for the average passenger time spent in the metro-line (Y_1), and for the fullness rate of carriages (Y_2) responses are respectively 606 seconds and 50,2 percent. After optimization study the fullness rate of carriages (Y_2) is decreased by 0.7 % and the average passenger time spent in the metro-line (Y_1) is decreased by 7 seconds (1.14 %) respect to current values.

Next, an optimum point is found for *STP*. For $X_1 = 10.0, X_2 = 5.5, X_3 = 10.0, X_4 = 15.0$ factor levels the obtained values for the average passenger time spent in the metro-line (Y_1), and for the fullness rate of carriages (Y_2) responses are respectively 590 seconds and 49.9 percent. After optimization study the fullness rate of carriages (Y_2) is decreased by 9.9 % and the average passenger time spent in the metro-line (Y_1) is decreased by 42 seconds (6.65 %) respect to current values.

Last, an optimum point is found for *SNP*. For X_1 = 15.0, X_2 = 14.5, X_3 = 15.0 factor levels the obtained values for the average passenger time spent in the metro-line (Y_1), and for the fullness rate of carriages (Y_2) responses are respectively 799 seconds and 50.0 percent. After optimization study the fullness rate of carriages (Y_2) is increased by 10.7 % and the average passenger time spent in the metro-line (Y_1) is increased by 105 seconds (15.13 %). Although the current average passenger time spent in the metro-line (Y_1) value is less than optimum result, it violates the fullness rate of carriages requirement (50%). The current and proposed headways and responses are shown in Table 6.1

Table 6.1 Current and proposed headways and responses

Problem	Current							Proposed						
	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2	X_1	X_2	X_3	X_4	X_5	Y_1	Y_2
WMP	10,0	7,5	5,0	7,5	10,0	613	46,2	9,8	9,9	7,1	9,0	7,8	625	50,3
								10,0	10,0	7,5	7,8	7,7		
WAP	7,5	5,0	7,5	10,0	15,0	613	50,9	7,3	6,2	6,8	10,0	10,5	606	50,2
STP	10,0	7,5	10,0	15,0	X	632	59,8	10,0	5,5	10,0	15,0	X	590	49,9
SNP	15,0	10,0	15,0	X	X	694	39,3	15,0	14,5	15,0	X	X	799	50,0

The current and proposed train numbers are shown in Table 6.2. Train numbers are calculated by using formula (4.1).

Table 6.2 Current (*C*) and proposed (*P*) train numbers

Problem	X_1		X_2		X_3		X_4		X_5	
	C	*P*	*C*	*P*	*C*	*P*	*C*	*P*	*C*	*P*
WMP	4	5	6	4	8	6	6	5	4	6
		4		4		6		6		6
WAP	6	6	8	7	6	6	4	4	3	4
STP	4	4	6	8	4	4	3	3	X	X
SNP	3	3	4	3	3	3	X	X	X	X

The main contribution of the current study is to (1) solve the optimization problem confronted in Izmir Metro Company by using *RSM*. To the best of our knowledge this is the first study using *RSM* for a public urban transport system, and (2) provide direction for public transport providers to find out optimum solutions using *RSM*.

REFERENCES

Angün, E, Kleijnen, J.P.C, Hertog, D.D., & Gürkan, G. (2002). Response Surface Methodology Revisited. In Proceedings of Winter Simulation Conference, 377-383.

Aytuğ, H., Doğan, C.A., & Bezmez, G. (1996). Determining the Number of Kanbans: A Simulation Metamodeling Approach, Simulation, 67:1, 23-32.

Barton, R.R. (1994). Metamodeling: A State of the Art Review. In Proceedings of Winter Simulation Conference, 237-244.

Batmaz, I., & Tunali, S. (2003). Small Response Surface Designs for Metamodel Estimation. European Journal of Operational Research, 145, 455-470.

Box, G.E.P., & Draper, N.R. (1987). Empirical Model-Building and Response Surfaces. New York: McGraw-Hill, Inc.

Brightman, H.J. (1978). Optimization Through Experimentation: Applying Response Surface Methodology, Decision Sciences, 9, 481-495.

Carey, M., & Carville, S. (2003). Scheduling and Platforming Trains at Busy Complex Stations. Transportation Research Part A, 37, 195-224.

Castillo, E.D., Montgomery, D.C., & McCarvile, D.R. (1996). Modified Desirability Functions for Multiple Response Optimization. Journal of Quality Technology, 28, 337-345.

Chiang, T.W., Hau, H.Y., Chiang, H.M., Ko, S.Y., & Hsieh, C.H. (1998). Knowledge-Based System for Railway Scheduling. Data & Knowledge Engineering, 27, 289-312.

Cordeau, J.F., Toth, P., & Vigo, D. (1998). A Survey of Optimization Models for Train Routing and Scheduling. Transportation Science, 32, 380-404.

D'Angelo, A., Gastaldi, M., & Levialdi, N. (1998). Performance Analysis of A Flexible Manufacturing System: A Statistical Approach. International Journal of Production Economics, 56-57, 47-59.

Dabbas, R.M., Chen, H.N., Fowler, J.W., & Shunk, D. (2001). A Combined Dispatching Criteria Approach to Scheduling Semiconductor Manufacturing Systems. Computers & Industrial Engineering, 39, 307-324.

Derringer, G.C. (1994). A Balancing Act: Optimizing A Product's Properties. Quality Progress, 27, 51-58.

Derringer, G.C., & Suich, R. (1980). Simultaneous Optimization of Several Response Variables. Journal of Quality Technology, 12, 214-219.

Dorfman, M.J., & Medanic, J. (2003). Scheduling Trains on A Railway Network Using A Discrete Event Model of Railway Traffic. Transportation Research Part B, 37, 81-98.

Durmisevic, S., & Sariyildiz, S. (2001). A Systematic Quality Assessment of Underground Spaces-Public Transport Stations. Cities, 18, 13-23.

Ferrari, P. (1999). A Model of Urban Transport Management. Transportation Research Part B, 33, 43-61.

Friedman, L.W. (1995). The Simulation Metamodel. Massachusets: Kluwer Academic Publishers.

Fuller, D., & Scherer, W. (1998). The Desirability Function: Underlying Assumptions and Application Implications. IEEE International Conference, 4, 4016-4021.

Gharbi, A. & Kenne, J.P. (2000). Production and Preventive Maintenance Rates Control for a Manufacturing System: An Experimental Design Approach. International Journal of Production Economics, 65, 275-287.

Groenendaal, W.J.H.V., & Kleijnen, J.P.C. (1996). Regression metamodels and design of experiments. In Proceedings of Winter Simulation Conference, 1433-1439.

Hood, S.J., & Welch, P.D. (1993). Response Surface Methodology and Its Application in Simulation. In Proceedings of Winter Simulation Conference, 115-122.

Horng, H.C., & Cochran, J.K. (2001). Project Surface Regions: A Decision Support Methodology for Multitasking Workers Assignment in JIT Systems. Computers & Industrial Engineering, 39, 159-171.

Huisman, T., Boucherie, R.J., & Dijk, N.M.V. (2002). A Solvable Queueing Network Model for Railway Networks and Its Validation and Applications for the Netherlands. European Journal of Operational Research, 142, 30-51.

Irizarry, M.D.L.A., Wilson, J.R., & Trevino, J. (2001a). A Flexible Simulation Tool for Manufacturing-Cell Design, I: Model Structure, Operation, and Case Study. IIE Transactions, 33, 827-836.

Irizarry, M.D.L.A., Wilson, J.R., & Trevino, J. (2001b). A Flexible Simulation Tool for Manufacturing-Cell Design, II: Response Surface Analysis and Case Study. IIE Transactions, 33, 837-846.

Jothishankar, M.C., & Wang, H.P. (1993). Metamodelling a Just-in-time Kanban System. International Journal of Operations & Production Management, 13, 18-36.

Kleijnen, J.P.C., & Sargent, R.G. (2000). Theory and Methodology-A Methodology for Fitting and Validating Metamodels in Simulation. European Journal of Operational Research, 120, 14-29.

Kreuger, P., Carlsson, M., Olsson, J., Sjöland, T, & Aström E. (1997). Trip Scheduling on Single Truck Networks-The TUFF Train Scheduler. In CP97 Workshop on Industrial Constraint-Directed Scheduling, 1-12.

Law, A.M., & Kelton, W.D. (1991). Simulation Modelling & Analysis. (2nd ed.). Singapore: McGraw-Hill, Inc.

Leon, N.A., & Cabrera, F.M. (1997). Improvement of Igniters' Quality Characteristics Using Experimental Design. Computers & Industrial Engineering, 33, 145-148.

Li, J.P. (2000). Train Station Passenger Flow Study. In Proceedings of Winter Simulation Conference, 1173-1176.

Mackett, R.L., & Edwards, M. (1998). The Impact of New Urban Public Transport Systems: Will the Expectations Be Met? Transportation Research Part A, 32, 231-245.

Madu, C.N. (1990). Simulation In Manufacturing: A Regression Metamodel Approach. Computers & Industrial Engineering, 18, 381-389.

Mahadevan, B., & Narendran, T.T. (1993). Theory and Methodology-Buffer Levels and Choice of Material Handling Device in Flexible Manufacturing Systems. European Journal of Operational Research, 69, 166-176.

McHaney, R.W., & Douglas, D.E. (1997). Multivariate Regression Metamodel: A DSS Application in Industry. Decision Support Systems, 19, 43-52.

Mollaghasemi, M., Evans, G.W., & Biles, W.E. (1991). An Approach for Optimizing Multiresponse Simulation Models. Computers & Industrial Engineering, 21, 201-203.

Montgomery, D.C. (2001). Design and Analysis of Experiments. (5th ed.). New York: John Wiley & Sons, Inc.

Montgomery, D.C., & Runger, G.C. (1999). Applied Statistics and Probability for Engineeers. (2nd ed.). New York: John Wiley & Sons, Inc.

Myers, R.H., & Montgomery, D.C. (1995). Response Surface Methodology: Process and Product Optimization Using Designed Experiments. New York: John Wiley & Sons, Inc.

Myers, R.H., Montgomery, D.C., Vining, G.G., Borror, C.M., & Kowalski, S.M. (2004). Response Surface Methodology: A Retrospective and Literature Survey. Journal of Quality Technology, 36, 53-77.

Neter, J., Kutner, M.H., Nachtsheim, C.J., & Wasserman, W. (1996). Applied Linear Statistical Models. (4th ed.). Boston: McGraw-Hill, Inc.

Osborne, D.M., Armacost, R.L., & Edwards, J.P. (1997). State of the Art in Multiple Response Surface Methodology. IEEE International Conference, 4, 3833-3838.

Parkes, D.C., & Ungar L.H. (2001). An Auction-Based Method for Decentralized Train Scheduling. In Proceedings of the Fifth International Conference on Autonomous Agents, 43-50.

Pegden, C.D., Shannon, R.E., & Sadowski R.P. (1990). Introduction to Simulation Using SIMAN. New York: McGraw-Hill, Inc.

Priemus, H., & Konings, R. (2001). Light Rail in Urban Regions: What Dutch Policymakers Could Learn From Experiences in France, Germany and Japan. Journal of Transport Geography, 9, 187-198.

Ribardo, C., & Allen, T.T. (2003). An Alternative Desirability Function for Achieving "Six Sigma" Quality. Quality and Reliability Engineering International, 19, 227-240.

Santos, M.I.R, & Nova, A.M.O.P. (1999). The Main Issues in Nonlinear Simulation Metamodel Estimation. In Proceedings of Winter Simulation Conference, 502-509.

Sargent, R.G. (1991). Research Issues in Metamodeling. In Proceedings of Winter Simulation Conference, 888-893.

Sinclair, M., & Oudheusden, D.L.V. (1997). Case Study-In Heavily Congested Cities. European Journal of Operational Research, 103, 18-27.

Sridharan, R., & Babu, A.S. (1998). Multi-Level Scheduling Decisions in A Class of FMS Using Simulation Based Metamodels. Journal of the Operational Research Society, 49, 591-602.

Şahin, İ. (1999). Railway Traffic Control and Train Scheduling Based on Inter-Train Conflict Management. Transportation Research Part B, 33, 511-534.

Tunali, S, & Batmaz, I. (2003). A Metamodeling Methodology Involving Both Qualitative and Quantitative Input Factors. European Journal of Operational Research, 150, 437-450.

Wu, C.F.J., & Ding, Y. (1998). Construction of Response Surface Designs for Qualitative and Quantitative Factors. Journal of Statistical Planning and Inference, 71, 331-348.

Wu, J., & Hounsell, N. (1998). Bus Priority Using Pre-Signals. Transportation Research Part A, 32, 563-583.

Yeh, C.H., Deng, H., & Chang, Y.H. (2000). Perspectives for Practice-Fuzzy Multicriteria Analysis for Performance Evaluation of Bus Companies. European Journal of Operational Research, 126, 459-473.

Yildiz, G. (2003). Simulation Optimization of Dual Resource Constrained CONWIP Controlled Lines Using Response Surface Methodology, PhD Thesis, Dokuz Eylül University, Department of Industrial Engineering, Izmir, Turkey.

Yu, B., & Popplewell, K. (1994). Metamodels in Manufacturing: A Review. International Journal of Production Research, Vol. 32, 787-796.

WEB_1. (2004). NIST/SEMATECH e-Handbook of Statistical Methods' web site. http://www.itl.nist.gov/div898/handbook/, 01/05/2004.