A RISK-AVERSE APPROACH
FOR SETTING TRANSIT LINE FREQUENCIES

(TOPLU TAŞIMA SEFER SIKLIKLARINI BELİRLEMEDE
RISKE DUYARLI BİR YAKLAŞIM)

by

İlayda ÜLKÜ, B.S.

Thesis

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE
in
INDUSTRIAL ENGINEERING
in the
INSTITUTE OF SCIENCE AND ENGINEERING
of
GALATASARAY UNIVERSITY

January 2014

A RISK-AVERSE APPROACH
FOR SETTING TRANSIT LINE FREQUENCIES

(TOPLU TAŞIMA SEFER SIKLIKLARINI BELİRLEMEDE
RISKE DUYARLI BİR YAKLAŞIM)

by

İlayda ÜLKÜ, B.S.

Thesis

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE

Date of Submission : Jan 03, 2014
Date of Defense Examination: Jan 17, 2014

Supervisor : Prof. Dr. Orhan FEYZIOGLU
Committee Members : Yrd. Doç. Dr. Zeynep ŞENER
Yrd. Doç. Dr. Gülfem ALPTEKİN

Acknowledgements

In the first place, I would like to express my gratitude to all those who gave me the possibility to complete this thesis. I am deeply indebted to my supervisor Prof. Dr. Orhan FEYZIOGLU for his help, supervision and guidance from the early stage of this thesis. He spent his valuable time for me and supported me with his wide knowledge. Once again gratefully thanks to him for sharing his documents and experience with offering suggestions for improvement and writing of this thesis to make this study finished accurately on time. I would also like to thank to my family for their support, concern, patience and love.

İlayda ÜLKÜ
İstanbul, 03/01/2014

Table of Contents

List of Figures

List of Tables

Abstract

The traditional objective to transit network frequency setting is the minimization of total in-vehicle and station waiting times. The data used in this decision-making process are generally the mean forecasted travel times and travel demand. When the bus line frequencies are set based on these data, passengers occasionally experience excessive travel times due to the changing travel network conditions and congestion. In this study, we obtain optimum line frequencies by considering stochastic travel times and demand, and use the conditional-value-at-risk measure to control the possible large realizations of random outcomes. We characterize the random network parameters by a finite set of scenarios and propose a risk-averse mathematical model. In this bilevel model, the network authority's objective is to increase the system reliability whereas passengers' objective is to decrease their expected travel time. To solve the model, we propose a method that integrates mathematical programming approaches with the genetic algorithm. The validity of the modeling approach is shown with a numerical study.

Résumé

L'objectif traditionnel pour le problème d'ajustement des fréquences de passage dans les réseaux de transport public est la minimisation du temps total passé en route et dans les stations. Les données utilisées pour ce processus de décision sont généralement les valeurs probables des durées et des demandes de voyage. Quand les fréquences des lignes sont ajustées selon ces données, les passagers prennent occasionnellement des temps excessifs pour voyager à cause des conditions dynamiques du réseau et des embouteillages. Dans cette étude, nous identifions les fréquences optimales des lignes en considérant les durées et les demandes de voyage stochastiques, et nous utilisons la mesure valeur à risque conditionnelle afin de contrôler les réalisations à grande échelle des résultats aléatoires. Nous caractérisons les paramètres aléatoires du réseau avec un ensemble dénombrable de scenarios et nous proposons un modèle mathématique visant à éviter les risques. Dans ce modèle à deux niveaux, l'objectif de l'administrateur du réseau est d'améliorer la fiabilité du réseau tandis que l'objectif des passagers est de réduire leurs temps de voyage prévus. Pour résoudre ce problème, nous proposons une méthode qui intègre les approches de programmation mathématiques avec l'algorithme génétique. La validité de notre approche de modélisation est démontrée avec une analyse numérique.

Özet

Toplu taşıma sistemlerinde hat sefer sıklıkları belirlenirken geleneksel yaklaşım toplam araç içi yolculuk ve durakta bekleme sürelerinin en aza indirilmesidir. Ancak uygulamada gerek yolculuk süreleri gerekse talepler için yaklaşık tahmini değerler kullanılmaktadır. Bu verilere dayanılarak belirlenen hat sefer sıklıkları, özellikle trafik sıkışıklığının değişkenlik göstermesi ile yolcuların zaman zaman aşırı sürelerde yolculuk etmelerine neden olmaktadır. Bu çalışmamızda, eniyi hat sefer sıklıkları belirsiz ağ koşullarından doğan stokastik yolculuk süreleri ve talepleri dikkate alınarak belirlenmiştir. Risk ölçütü olarak koşullu riske maruz değer (CVaR) kullanılmıştır. Temel olarak CVaR, yolculuk süre ve taleplerine bağlı rassal çıktıların olası büyük değerler almasını kontrol etmek amacıyla kullanılmaktadır. Rassal parametreler sonlu sayıda senaryolar ile temsil edilmiş ve riske duyarlı bir matematiksel model kurulmuştur. İki seviyeli olan bu modelde, ulaşım yönetiminin amacı bütün sistemin güvenilirliğini arttırmak iken yolcuların amacı beklenen yolculuk sürelerini en aza indirmektir. Çalışmada matematiksel programlama yaklaşımlarını genetik algoritma ile beraber kullanan bir çözüm yöntemi önerilmiştir. Modelleme yaklaşımının geçerliliği yürütülen sayısal çalışma ile gösterilmiştir.

1 INTRODUCTION

With the advances in the technology and industry, the migration from the rural areas brings rapid urbanization. As urbanization is not well planned in less developed countries, many problems occur. The transportation problem is one of the most important among them.

Transportation problem is also growing day by day in Turkey, due to the rapid increase in urbanization and industrialization, and the increase in the cities' population for various reasons. In the face of increasing mobility on the roads, it is difficult for people to go from one place to another. According to the Confederation of Employers' Unions of Turkey (TİSK) report, the average rate of population growth between 2004-2020 for some countries will be as follows: "India (1.3), Turkey (1.2), Ireland (1.2), Mexico (1.1), Brazil (1.1), Australia (0.9), U.S. (0.9), Canada (0.8), China (0.6) ". Table 1 shows population growth in major developing countries.

Table 1.1: Population growth for major developing countries

	Population (million people)		
	2003	**2008**	**2012**
Turkey	70.8	71.5	75.6
Brazil	179	191.9	204
India	1076.3	1158	1205
China	1292.3	1328	1343.2

Gross Domestic Product per capita (GDP) is often considered as an indicator of a country's standard of living. As the purchasing power increases, the national wealth also rises. According to the Automotive Industry Association report, 12 percent rise for the total automotive market in Turkey is monitored comparing the periods of January-June 2012 and 2013. With respect to the Turkish Statistical Institute (TUIK) data, the number of registered vehicles in traffic by the end of April 2013 has increased by 84,921 vehicles compared to the previous month, and reached to 17,350,448. Increasing car ownership in the societies causes many important problems. These problems can be partially solved by public transportation systems so they became one of the most important issues in the cities.

With the growing population, especially in the developing countries, the increase in waiting times at stops and in-vehicle traveling times affect the lives of people in a negative way. According to the report “Access to transport for the urban poor in Asia” (UN-HABITAT, 2009) the waiting time of passengers at bus stops is between 15 to 30 minutes in Bangalore, India. Network disruptions due to the factors such as bad weather, accidents, congestion and vehicle breakdowns additional increase the total traveling times of passengers. People become more irritated and show less performance in work and educational activities. In other words, unreliability of transit systems has direct and indirect social and economic impacts.

An important goal in urban transportation is to improve the service quality of passengers. One dimension of the service quality is service reliability. For this reason, unexpected situations should be taken into account during the planning process. However, the opening, frequency setting and timetabling of transit lines is mostly based on the average or expected values of traveling times and demand. This approach sometimes causes huge delays in the system for rare but probably occurring situations. In this study, we develop a risk-averse mathematical programming model to limit the large costs of the worst case scenarios. The model focuses only on the frequency setting stage in transit planning. In-vehicle travel times and travel demands are assumed to be stochastic, and their various probable realizations are included in the model as scenarios. The hierarchy in the decision-making or the relationship between the network manager and the network users is reflected to the model with a bilevel structure. To model the risk-averse behavior of the network manager, conditional value-at-risk measure is used to decide on the total costs of the scenarios. To our knowledge, this is a novel approach for the planning of transit systems by means of mathematical programming.

The thesis is organized as follows. In Chapter 2, we provide a review of previous works done and our motivation to develop the risk-averse transit network frequency setting model. In Chapter 3, we first give the necessary background on risk measures, bilevel programming and transit route choice models, and then present the developed mathematical model. The details of the proposed solution methodology for the transit frequency setting problem are introduced in Chapter 4. Chapter 5 contains numerical results obtained on an illustrative example. Finally, Chapter 6 includes concluding remarks and perspectives.

2 LITERATURE SURVEY

Generally, with the beginning of designing a network to schedule the buses to lines is very important for planners. On the other hand, to satisfy passengers willing to get on the bus is another critical topic. To get along with, a very important transit planning process is needed. Therefore, transit process has become a popular area for lots of researchers. In the literature, there are a wide range of studies about transit line frequency setting problems. The main aim of these studies is to find reasonable solutions for both planners and passengers. To cover this problem, there are some processes to be done. First of all, these processes begin with the design of the network which is related to find an appropriate network according to demand. In addition, this demand includes passengers who are willing to go from the origin to the destination point. The next important process is frequency setting of the lines in the network to decide on available frequencies of the buses. Another critical process is timetabling of lines in the network where the departure times of buses along the line is determined.

2.1 Transit Planning Process

Guihaire and Hao (2008) introduce a global review of transit network design and scheduling. The study includes the goals of strategic and tactical transit planning. They establish a terminology proposal in order to name sub-problems and thereby structure the review. The public transit planning has five steps: process route design, frequency setting, timetabling, vehicle scheduling, and crew scheduling.

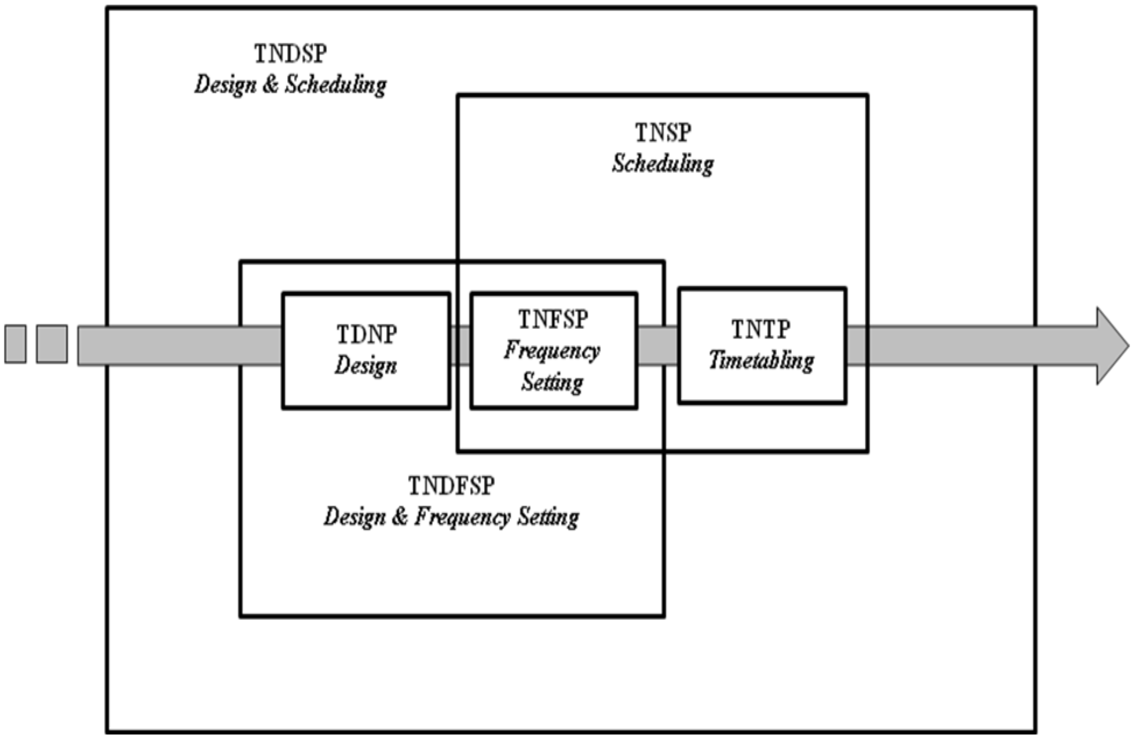

Figure 2.1: The structure of transit network problems (TNP)

2.1.1. Transit Network Design

For the transit network design problem several features exist and they can be considered both in the objective and the constraints. Therefore, we only list them here as a unique set of features.

Area coverage: The ratio of the estimated demand that can be served by public transit can be computed in several ways (Spasovic et al., 1993). However, it is well known that this ratio can be affected by route length, density, bus stop and route spacing (Murray, 2003). As for example, people living within 400–500 m from a bus stop are part of this ratio.

Existing network: As sometimes it might be undesirable to disrupt service on already existing lines, it can be mandatory to consider the existing network.

Demand satisfaction: When bus stops are too distant from users' origin or destination or when the trip takes too long, the demand can be considered unsatisfied. This is of course a crucial issue. Demand satisfaction also requires to go through a transit trip assignment process. As for example, the user will switch to another means of transportation if the trip requires more than two transfers.

Route and trip directness: To design a bus network enabling users to travel as directly as possible from their origin to their destination and to walk the shortest distance to reach the first and final bus stop is an important issue. Hence, limits can be imposed on the distance that one user can cover in the transit network. Directness can depend on the route's deviation from a linear or shortest path. To compute trip directness for each user, it is necessary to go through a passenger trip assignment process having objective such as shortest path or smallest number of transfers (Desaulniers and Hickman, 2007).

Number of lines or total route length: A general objective of the operator is to minimize the total route length in the perspective of reducing the number of vehicle and crew resources needed to sustain the global transit system. The number of lines can alternately be considered. Moreover, routes should neither be too short nor too long for profitability reasons. In a general manner, if a trip requires more than two transfers, it is assumed that the user will switch to another means of transportation.

2.1.2. Transit Network Frequency Setting

The demand may vary according to the day of the week, time of the day, time of the year. To operate an efficient network and provide a satisfying service, demand variation should be taken into consideration. Collecting this data is a very complex task and the transit agency has a great responsibility. The origin-destination (OD) matrices are used in the frequency setting step. The following issues should be considered while formulating the transit network frequency setting problem.
Demand satisfaction: To avoid overcrowding and excessively large headways and also, to reduce waiting and transfer times, lines frequencies should match the demand.

Number of line runs: The problem to find optimal number of runs for each line has a multiobjective nature. While users want to use wide range area of the network, the operators desire to minimize the number of resources. Also, line frequencies depend on the available fleet size and the capacities of the buses. Bus running times are associated to each route of the network for the frequency setting problems.

2.1.3. Transit Network Timetabling

Timetable includes departure times from all the stops served by each line run in the network. The expected arrival time at the final stop and the expected departure times from each bus stop on the route defines each line run's timetable. To compute the timetables, running times are necessary. The time coverage of the line is defined in the transit network frequency setting step by determining the lines frequencies. The minimization of passengers waiting times is obtained by the level of each transfer for a better service quality. Each transfer should be deduced from the OD matrices. Some of the features included in the transit network timetabling problem are listed below.

Demand satisfaction: To enhance the passengers' mobility, passengers travel time should be minimized.

Fleet size: The line runs of the transit network timetable gives the vehicle schedules.

Transfer coordination: Each transfer zone and associated lines should be included in the global network timetable (Guihaire & Hao 2008).

2.1.4. Vehicle Scheduling

A feasible sequence of line runs is obtained in the vehicle scheduling step. An optimal schedule should minimize fleet size and operational cost. The objective is to find an assignment of trips to vehicles, such as each vehicle performs a feasible sequence of trips and every trip should cover once.

2.1.5. Crew Scheduling

In this step drivers are assigned to the bus services. Every driver finishes the timetable with a given period. Maximum number of consecutive working days are included as a constraint while assigning the drivers to the buses.

2.2 Transit Frequency Setting Problem

All of the processes that we have defined in Section 2.1 can be used in the transit planning. In this study, we will only focus on the transit frequency setting problem. The transit network frequencies setting problem arise from choosing sufficient frequencies for lines on the network. An ideal frequency setting should satisfy the request for both the passengers and travel agency. The decrease in the operator's costs, and in the fleet size is also included in the frequency setting problem (Constantin & Florian 1995; Guihaire & Hao 2008). There are studies related with the capacitated models, seat capacitated models, and uncapacitated models. In these studies, the minimization of total travel time is associated with the expected cost.

In their pioneering study, Constantin and Florian (1995) propose a model that aims to find the optimum line frequencies minimizing the total of expected travel and waiting time of the users on the network while considering fleet size constraints and bounds on the frequencies. They first formulate the problem as a mixed integer nonlinear problem but they then reformulate it as a min-min nonlinear bilevel program. This latter model is solved with a gradient projection type algorithm where a subgradient is obtained at each iteration by solving the lower level transit assignment problem. Computational results are provided for the transit networks of the cities of Stockholm, Winnipeg and Portland.

Cominetti & Correa (2001) consider the common-line problem under congestion, passenger assignment and equilibrium in the network. The authors developed the model by using the wardrop equilibrium model. The model includes effects of congestion over the passengers' decisions. Assignment of the passengers to bus lines is based on the common-line frequencies. The main objective of the study is to consider a network equilibrium model with congestion. The objective function is affected by the flow distribution, the waiting times and travel times according to the congestion. The model includes an assumption with the passengers' decision to travel with the shortest paths. Also, the authors accept that each passenger can easily get on the first arriving bus. Nevertheless, boarding to the bus which come first, avoid to increase the passenger’s waiting time. Because of the limited bus capacity, the model assumes that the line frequencies are decreasing functions of the lines. Therefore, according to the authors, the lines are considered as a differentiable effective frequency function. In the literature the researchers take into account different types of arcs when they are modeling general transit networks. These arcs are: boarding, alight, on-board, and walk arcs. While the other arcs correspond to services which are always available, the waiting processes affect mainly the boarding arcs. Thus, the waiting times are either zero or negligible. Therefore, the model is considered as an infinite frequency for the situation. When the transit time and boarding probabilities are considered, then the infinite frequencies are replaced with a constant frequency for one or more lines. Also, time limit is considered. Cominetti & Correa (2001) also consider the passenger assignment and equilibrium in general networks. The common-line scheme is used in the model with a dynamic programming approach. Moreover, the equilibrium existence and uniqueness conditions are included which shows a simple convex function, and gives the circumstances for the model. Besides, the occurrence of flow in the common lines is represented as some unexpected property. Thus, by restricting the user’s choices, increasing the transit time can be achieved. The authors have pointed out that the system performance is not influenced by the flow increment.

In other respects, there are a variety of studies where transit networks with capacitated vehicles are considered. One of them is due to Lam et al. (2002) that considers the capacity constrained transit assignment problem with elastic line frequency. The line frequency is related to the passenger flows on transit lines. A fixed point problem with the line frequency is used, when the passenger waiting time and the line capacity are dependent on the line frequency. The model is developed by using a stochastic user equilibrium transit assignment model. It also includes both congestion and elastic line frequency. Besides, not only line capacity but also passenger waiting time is related with the line frequencies. To observe the effects of the suggested model and the effectiveness of the proposed solution algorithm, an illustrative example is considered. The amount of time that is spent in the intermediate stations on the transit line is considered as the effects of the model. The path between an origin and a destination (OD) is defined as a sequence of transfer stops that is the transit route. The authors consider the frequency of each transit line to be dependent on the vehicle dwelling time at stations, with the assumptions of fixed transit fleet size and constant in-vehicle travel time. The line frequency is related to the passenger flows on transit line. A stochastic user equilibrium transit assignment model with congestion and elastic line frequency is considered with the formulated mathematical programming problem.

Florian (2003) investigates transit line frequency setting in three different problems: linear cost model, nonlinear cost model, and variable frequency model. When the travel times and the frequencies are constant, the model is a linear cost model. In the linear cost model, the distribution of the times of arriving vehicles to the destinations is given for each line. The objective is to minimize the sum of expected waiting and the travel time. The author assumes that the network has four types of arcs: wait arcs, in vehicle, alighting, and walk arcs. In the first part of the solution algorithm, not only from the destination nodes to all origins the arc flow is calculated, but also, the expected travel times are computed for each node. In the second part, from all origins to the destination, the demand is distributed to the arcs. The author mentions that the algorithm is applied for each destination. When the travel time is not constant, but an increasing function then the model is the nonlinear cost model which is convex. The passengers choose strategies which minimize the expected travel times. Since, linear cost model can be separated, the nonlinear problems can not be separated. This problem can be solved with convex cost optimization problem. Moreover, the stopping criterion can be used as maximum number of iterations. According to the research, all passengers who board the bus first have a better chance to find a seat. Therefore, the variable frequency or capacitated model is also discussed. If the transit network is crowded and the passengers can not board to the first vehicle, the effective frequency is developed. The linear cost model has been applied in various software packages. As a result of the study, the variable frequency model is the most effective model for congested systems.

The effects of congestion on the expected waiting and travel times are investigated in Correa and Cominetti (2001) as a frequency based route choice model for congested transit networks. The authors formulate a new characterization of the equilibrium, and this formulation for the equivalent optimization problem does not include the gap function. The proposed solution algorithm can be then applicable on the large scale networks. Path based transit network equilibrium models and the flow with travel times is included in the model formulation. The objective of the study can be explained as to define an optional formulation for the frequency based transit equilibrium model correctly. Cepeda et al. (2006) have enhanced the same model by adding the capacity constraints. Many arcs have no waiting time and these arcs are assigned with the infinite frequencies, when modeling general transit networks. In the model's formulation, it is sufficient to consider the convention in all calculations by replacing the infinite frequencies and then these frequencies are converted to number one. Especially, when getting on the route of the network, the arcs include infinite frequencies with travelling time of the average of the transfer time from origin to the destination node. A heuristic minimization method is used with the value of the gap function to determine the optimality difference of the computed solutions. The easiest way to solve equilibrium problems is a well known method of successive averages that is used in transportation applications. By fixing the travel times and the frequencies at the values determined by the current flows, the method computes a transit network equilibrium for the linear network at each iteration. Then these flows are updated by averaging the previous iterate solution. To solve the linear program can found the equilibrium of the linear cost network. This linear program calculate the shortest path for each destination. Then, the linear program defines the necessary arc-destination flow vector. This paper introduce a different view then the previous studies on transit assignment model in congested networks.

With the improvement of search algorithms, important heuristic approaches is done in the literature. Tom and Mohan (2003) apply genetic algorithms in their study. The authors define a group of routes and also finds the correlated frequencies which are relevant with transit route network design. They have formulated a problem which reduces the overall cost with the operational constraints. This cost determines overall travel time of passengers, and operating cost. The authors determine a solution process in two phase. In the first phase, a group of route is defined to select a route through this group by applying a candidate route generation algorithm. The second phase includes a necessary route set that is chosen as a solution from the candidate routes which was found in the first phase by using genetic algorithms. In the model, the frequency coded model and the route is determined at the same time. A medium sized network is created to see the performance of the model of the route network design.

Bus transit frequency design with the capacity constraints is also examined by Monyrath (2006) with variable demand which is defined under fixed total demand. The study also presents the mutual effect of different type of vehicles in the network traffic, such as cars, buses and motorbikes are taken into consideration. Designing bus line frequency is not preferred for this study. The authors used a combined mode split assignment model to determine transit demand. This assignment model is called modal split. Also, the flow of the mixed traffic's line is determined according to bus frequency. The study use "diagonalization algorithm" to provide the movement of the bus network and the mixed traffic network. The diagonalization algorithm is related to solve a series of standard user equilibrium programs repetitively. To reduce the overall travel cost of multi-modal network users is the objective of the study. According to the results, with using different initial frequencies, the optimal frequencies can approach to the global optimum. In the sensitivity analysis part, the increment in travel demand and the change in the operation cost is managed with the continuous modification of the transit line frequency.

Different than the above studies Teklu (2007) considers Monte Carlo simulation based stochastic process for frequency-based transit assignment. The author have pointed out to forecast the mean route flows and costs together by using the stochastic process model. A stochastic process model for transit assignment is considered based on realistic route cost estimates which provides the forecasts of mean route flows and costs together. A Monte Carlo simulation approach is adopted to model for the capacity constraints. A random utility model for route choice is suggested, because of the variation on the passengers' reflection to the congestion. Therefore, the model includes a bus and passengers simulator, a random utility model for transit route choice, and a learning process model. Probit route choice model is developed to take into account the cost correlations to see the difference with alternative routes. Deterministic frequency-based transit assignment models assume all passengers have a knowledge of the alternative routes, the associated costs. By using a simple network, numerical experiments are used to illustrate model effectiveness. In addition, there is some extra considerations to include in the model. For instance, to find empty seats in the buses is important for passengers. But, finding empty seats include stochasticity in the model. Deterministic models for transit assignment do not provide information of variability which could help transport planners to make more informed decisions.

In their recent studies, Schmöcker et al. (2008) consider dynamic frequency-based transit assignment model. The objective of the model is to provide the minimum gab between schedule-based and frequency-based models. In highly congested networks a static frequency-based approach does not fit capacity problem efficiently. Because of the overcrowding some passengers can not board the first service arriving bus. Therefore, a ‘‘fail-to-board” probability is used to present the line capacity constraints. Also, the shortest hyperpath is affected by the fail-to-board probability in this line problem. According to the study, the passengers who could not board on the bus, are added to the demand of the following time interval. Therefore, this study is separated in lots of time intervals and the passengers route choice can be considered lots of time. In addition, the model is applied both in a small example network and in a case study in London. The case study which is applied in London is tested daily during the peak period of the transit capacity problems. Thus, the authors are applied a dynamic frequency based transit assignment model. Since, the problem is executed in London, the dynamic frequency based transit assignment models are convenient for such a crowded city. Because, a static approach can not solve this congestion problem. Moreover, because of crowdedness passengers could not get on the bus and they have to wait to board to the next bus, as in the case of congested cities. For that reason, congestion may continue for the next buses. The authors have some assumptions about the time intervals which are constant and the time intervals are selected as 15 min. Also, OD demand is changed during the dynamic approach implementation. In addition, different intervals of congestion is applied in the study. Hereby, congestion will affect passengers conclusion on the route. In the result of the study, passengers determine. Thus, passengers choose different exchange points when they are travelling in the off-peak.

Furthermore, another real case study that is considered with uncapacitated and capacitated transit networks is implemented in Santiago, Chile by Leiva et al. (2010). From the perspective of users and operators, in bus networks with large demand, finite stop services provide help for the passengers and the operators. The network's efficiency is necessary to determine the design of these services with frequency and vehicle size. To determine efficient frequency for the unplanned services on an urban bus corridor which means bus line is presented by minimizing social costs. The objective of the study is to select the line with sufficient frequency for each route section. The author's optimization approach includes minimizing overall costs that involve operator cost, waiting time of passenger, and in vehicle travel time. They assume that the travel demand is known. To decide on the minimum cost the authors follow some elements. To formulate a mathematical programming problem they need to determine these elements. First of all, the set of stops along the bus line need to be known. After, the distance between stops must be specified. In addition, travel demand between stops must be detected. Besides, the authors are used two different demand scenarios to decide on the good candidates for each service. Then, the model results give with the frequencies of the lines and necessary bus sizes. Also, the authors include the passengers' decision of transfer points and attractive lines. The passengers demand on each route can be fallowed by a set of variables which is presented by authors. The model is constrained with bus lines and vehicle capacity. With including these constraints, the model is solved as non linear integer model. After that the models are settled to decide on the type of vehicles and about the frequencies which should be used in the network. Moreover, to see the model and algorithm's affect a real bus line in

Santiago, Chile is handled, and also the paper introduce the solutions. According to the study, the solution algorithm achieves a satisfactory result on the real case problem.

In the same year, Yoo et al., (2010) introduce a method to model the transit frequency design problem with variable demand. According to the non-cooperative Stackelberg game, a bilevel optimization model is used. The authors use lots of assumptions in this model to investigate the frequency model. First of all, maximization of the demand for transit is dependent on one central which can be operator or government. Then, only two modes are used by passengers, such as, a private car or transit. In addition, there is constant overcrowding on the road network which means that the amount of transit flow which incurs some trouble on the network is similar with the utility of the private car. Moreover, all transit lines are statistically independent and exponential distribution is given for the inter-arrival times of the vehicles. Besides, passengers arrive randomly with a uniform distribution at every stop. Furthermore, in-vehicle travel time is constant on each route-section and it is specified with the level of overcrowding on the road network. The proposed model has upper level problem and also lower level problem. With the fleet size and frequency constraints a non-linear optimization model is formulated to maximize demand as the upper level problem. A gradient projection method is used to solve the upper-level model. A capacity constrained stochastic user equilibrium assignment model with variable demand is formulated for the lower level problem. Also, transfer delays is considered in the lower level problem. To determine the transfer delays of the upper level problem, the fixed line frequencies in the lower level problem is solved for each iteration for a given frequency. A new frequency is achieved by solving the upper level problem until the gap becomes enough small of two sequential frequencies found. Also, authors increase the total demand and discuss the results. At the same time, evaluation of transfer delays between transit lines to insure more realistic situation of real transit networks. Moreover, an iterative balancing method is modified for this study to solve the lower-level model. A small network is used to apply their model and algorithm. According to the results the algorithm is fitted to the optimal point.

One year later, Yu et al. (2011), have published bus frequency design with a bilevel programming model which finds the optimal bus frequencies. This study aims to decrease the passengers overall travelling time with the constraint on the total fleet size. Also, the authors take into account the users route choice behaviors. There is upper and lower level algorithms in the model. According to the optimal policy, the objective of lower level specify transit lines upon the network. Moreover, the effect of the vehicle congestion is presented among designing frequencies. On the other way, routes' bus frequencies are optimized related with the upper level of that is dependent on the passenger assignment as a nonlinear program. Furthermore, bilevel model is solved by using genetic algorithm and a label-marking method. Also, the upper level model deals with the frequency optimization problem which can be observed as a vehicle distribution problem of different routes. In this model the objective is to minimize total passenger travel time. There is limited number of vehicles. Therefore, the frequencies of routes have to satisfy the fleet size constrain. Because of this limited number of vehicles, some routes can occur congestion. Thus some passengers can change their way from the overcrowding to less crowded routes. The optimal frequency of each route is found, when the passenger distribution is near to equilibrium situations. As a comparison of the model and algorithms, two test examples are examined. According to the optimization results, the local service level can be improved. The bilevel formulation is introduced as a relation between passengers’ route choices which is decision takers and setting frequency which is a decision maker. First of all, initial frequency is defined for each route. Second part is the lower level model that the initial frequencies give the assignment of passengers to the network. Then, according to the number of passengers optimizing the frequency of each route which means origin destination path from the lower-level model, give the objective of the upper level model. Besides, again in the lower level model, with the determined frequencies from the upper level model passengers are assigned to the network again. These two models are used to optimize frequencies of routes and stops when a fixed solution for the frequencies and the passenger assignment on the routes is found. The first one shows the validity and feasibility of the bilevel model on the simple network. On the other hand in the second one, the real transit network in the city of Dalian, China is used to presents the performance of the model and the algorithms for optimizing frequencies as a solution approach.

In their recent work, Monyrath et al. (2011) consider an analytical model to decide the optimal frequency for urban bus transit by thinking about the interaction between different vehicles in urban mixed traffic with including variable demand. In-vehicle travel time impact the road congestion by reminding each mode to specifically considered in the formulation function of link performance. There are lots of objectives to think about the bus frequency optimization model to choose an optimal frequency of bus service. Therefore, the authors use total waiting time of bus passengers, total travel time of all passengers in the network, and the operating cost as the objective. On the transit network in mixed traffic and the passenger flow, the equilibrium flow of vehicle is computed by using multi-modal equilibrium assignment. Firstly, to find the optimal frequency of the bus transit the authors study on the changes of equilibrium link flows which are line frequencies. Also, each frequency assigned line is solved for the multi modal equilibrium assignment. As the solution algorithm first of all, a feasible line frequency vector is used which fits the constraint in the frequency design model. Moreover, on road network and on bus network, link flow vectors are found with modal split assignment model. Then, the objective function is calculated with this equilibrium link flow and the frequency which are included into the objective function. Moreover, this equilibrium is used according to the frequency which is selected by Hooke-Jeeves algorithm. The solution algorithm is repeated until better cost of the system which means better frequency that decreases the objective function. Besides, by repeating the processes the optimal frequency is found. to show the application of the model and algorithm, a numerical example is presented. Some significant decisions are emphasized on sensitivity analysis of the computational results. The proposed model deals with the problem from many points of view with the computational result of bus network design in mixed traffic.

A non-linear mixed integer programming model which involves the network route design and frequency setting problems simultaneously is introduced by Szeto and Wu (2011). In the investigated problem, the intent is to improve the existing bus services by reducing the number of transfers and the total travel time of the users. The objectives and the constraints of the problem are different from those in the previously existing literature. To solve the problem, a genetic algorithm is hybridized with a neighborhood search heuristic. The first algorithm deals with the route design problem while the second one tackles the frequency setting problem. A new solution representation scheme and specific genetic operators are developed to search all possible route structures. Compared with the current network design of Tin Shui Wai, a suburban residential area of Honk Kong, the proposed method can generate a design which can simultaneously reduce the number of transfers at least by 20.9% and the total travel time 22.7%.

Besides, Yu et al. (2011) use a parallel genetic algorithm to concern a study and apply their study in Dalian City, China about uncapacitated transit networks. The nonlinear program is formulated which includes the fleet size constraint. This study objectives to increase the quality of service and to decrease the cost of operations mainly. The passenger and operator costs are also introduced in the model. To determine the relative weights between passenger costs and operator costs, an integrated approach is also developed in the paper. Generally, when frequencies are too small, operators have to stand against operational costs. Meanwhile, when frequencies are too large, the service may not be sufficient, and the passengers can be unsatisfied. One key point of the study is a parallel genetic algorithm with a coarse-grained strategy and a local search algorithm based on Tabu search are considered to solve the headway optimization model. The authors solve the problem using real data which are collected from Dalian City, China to control the model applicability and the solution methodology. Results show that the reasonable resource assessment can increase the benefits of transit system.

Additionally, unlike the above study Gallo et al. (2011) focus in both uncapacitated and capacitated transit networks. The paper proposes some models and algorithms to solve the transit network design problem under the assumption of elastic demand. The authors deal with elastic demand in the transit network design problem, considering a regional metro network which determine the frequencies of a regional metro on a given network. The problem is used not only to increase the given network's efficiency but also to determine investments feasibility in the rail network. In this transportation system modal split is used: rail transit system, bus transit system, and private car system is included in the transportation system. To ignore the demand elasticity can not represent the actual objectives of the design. Thus, the elastic demand is taken into consideration. There are four objectives with assumption of elastic demand: operator costs, (rail and bus) transit user costs, car user costs, and external costs. The authors propose a heuristic solution algorithms where a Heuristic Local Search Algorithm is applied, and they use a meta-heuristic algorithm that is called as a Scatter Search. Another solution algorithm: Genetic Algorithm, is developed to compare the performance of the proposed algorithm. The authors develop a Genetic Algorithm to see a benchmark between these proposed algorithms. As a result of proposed algorithms, the comparison through scatter search, the Heuristic Local Search Algorithm and genetic algorithm, the scatter search gives better results in terms of objective function values. Because, the genetic algorithm generates solutions that are not necessarily local optima, while scatter search produces local optima. Thus, the genetic algorithm gives better solutions for a small network. On the other hand, the scatter search gives better solutions for a larger network. They prefer to test their algorithms and models not only in a small network but also in a real case network.

Similarly, to distinguish the trouble passengers who can take a seat in the vehicle, and who can not is handled as the schedule based equilibrium transit assignment model is observed by Hamdouch et al. (2011). Also, to allocate seats is interested in the model. The authors assume that some policies are used by the passengers, while they are travelling from the origin to the destination points. The authors develop an analytical model to take a seat in the stochastic nature of the passengers who take a seat and standing in the vehicle. Therefore, first come first serve principle is applied for the waiting passengers. They board to the first coming vehicle respectively. For example, the passengers who arrives to the station to wait the vehicle have priority to take a seat previously. Furthermore, users probabilities are assigned in an order to sit, to stand and failure to board by using dynamic network loading. The authors use their remaining travel time and time already spent in vehicle to assign the passengers in an order. Also, the standing passenger's motivation is depend up to the travel length and spending time on-board. If the time (travel length and spending time on-board) is long then the passenger's motivation increases to sit, vise versa. Moreover, the passengers may not get on the vehicle when it is full, they may wait for the vehicle to arrive. Unlike the Teklu's paper in 2007, the authors did not use the Monte–Carlo simulation about the seat availability. They prefer to use seat allocation model. The reason of this decision is explained as passengers' chance of getting a seat can be performed better in seat allocation model. A variational inequality is developed for the equilibrium conditions. Expected strategy costs as well as a vector-valued function can be involved in this equilibrium conditions. Therefore, successive averages method is created to find a solution. In each iteration of the solution strategies are generated to solve a dynamic program of the method. Computational solutions are given to see the impacts of the proposed algorithm about the travel policies and departure time due to the users' decisions.

Furthermore, there are also the transit networks studies that concentrate both the capacity constraints and the seat capacity constraints in the literature. Hadas & Shnaiderman (2012) present a stochastic demand and travel time that is used when uncertainty occurs in the transit line frequency setting problems. A supply chain optimization model where frequency setting is determined by collection data with stochastic properties, has two main cost elements to fix the optimization model. One of them is empty-seat driven, and the other is overload demand. The purpose of their problem is to reduce the vehicle capacity, and total cost. Deterministic data and stochastic demand is used to consider optimal frequency setting. According to the model the costs with considering stochastic demand, travelling with empty-seats and overcrowd, and travel time provides optimal frequency.

Finally, Codina (2012) introduce a paper about a variational inequality reformulation of a congested transit assignment model. The author extend the Cominetti and Correa's study (2001) to the to the common-lines problem for general multidestination networks under congestion. The reformulation of the congested transit equilibrium assignment model is performed by expressing the equivalent variational inequality. The case of strictly capacitated transit networks is explored under the scope of the authors new reformulation.

When we classify the literature studies; to solve the frequency setting problems, Constantin & Florian, 1995, Lam et al., 2002, Tom. & Mohan, 2003, Monyrath, 2006, Cepeda et al., 2006, Teklu, 2007, Schmöcker et al., 2008, Leiva et al., 2010, Yoo et al., 2010, Szeto & Wub, 2011, Yu, B. et al., 2011, Hamdouch et al., 2011, Gallo et al.,2011, Hadas & Shnaiderman, 2012, Cominetti et al., 2001 use Expected Cost models. Only Know, 2011 use Conditional Value at Risk (CVaR) in transportation problem. From the literature review one can see that Conditional Value at Risk (CVaR) is not used in the frequency setting problems until now.

3 MODEL FORMULATIONS

In this chapter, the details of the proposed transit network frequency setting model are presented. As this model in a based on the risk measure CVaR and has a bi-level structure, these concepts are first introduced.

3.1 Considering Risks in Transportation

Risk aversion is a perceptivity in psychology, economics, and finance, related with the behavior of people when they are faced with an uncertain situation to decrease that uncertainty. For instance, a risk-averse person can decide to put his or her money into a bank account with a low but assured interest rate, instead of deciding to put into a stock that can have higher expected returns, however this type of decision includes a chance of losing that amount of money. This is because this decision can lead to some loss of money, thus involves risk.

In the field of transportation, there are few studies that consider uncertainty and the associated risks in mathematical programming transport models. As for example, users' route choices in the presence of route costs uncertainty are investigated in Bell and Cassir (2002). Uncertainty about costs differs from variation in cost perception and the conventional approach is to add a safety margin based on the standard deviation of link cost. However, this requires the specification of the cost distribution. An alternative approach is presented in this study whereby the network user "plays through" all the possible eventualities before selecting his best route. A deterministic user equilibrium traffic assignment is shown to be equivalent to the mixed-strategy Nash equilibrium of an *n*-player, non-cooperative game. Then an *n*+*m* player, non-cooperative game is formulated, where *n* network users seek their best routes and *m* origin–destination specific demons penalize the network users maximally by failing links. The mixed-strategy Nash equilibrium of this game is shown to describe a risk-averse user equilibrium traffic assignment. A simple solution procedure is presented along with an illustrative example.

The main details of the literature considering capacity constraints are considered in Table 3.1. The table deals with the articles in the literature that includes capacity constraints, uncapacity constraints or seat capacity constraints. Also, the Table 3.1 is arranged according to years which includes main objective of the study, constraints and used method in the paper. Table 3.1 shows the papers with an application or a numerical example. Hence, the table incrementally increased according to years the change from year to year can be seen in transit network.

In transit frequency design problem, the most used objective function is minimizing overall cost that reduces time. Although, frequency is related with demand, there are lots of studies that consider fixed demand with the studies above from literature. In addition, frequency, load factor, and fleet size constraints are used as feasibility conditions in the transit frequency setting studies. With this thesis, we propose bi-level optimization model to obtain optimum line frequencies by considering stochastic travel times and demand with different type of scenarios. The overall model is solved by an adapted genetic algorithm.

Empirical studies have suggested that travel time reliability plays an important role in travelers' route choice behavior. In their study, Lo et al. (2006) develop an approach to relate the travel time variability due to stochastic network link capacity variations with travelers' risk aversive route choice behaviors. They postulate that travelers acquire the variability of route travel times based on their past experiences and factor such variability into their route choice consideration in the form of a travel time budget. This travel time budget varies with individuals and trip purposes and is related to the requirement on punctual arrivals. Moreover, all travelers want to minimize their travel time budgets. The authors formulate a multi-class mixed-equilibrium mathematical program to capture the route choice behaviors of travelers with heterogeneous risk aversions or requirements on punctual arrivals. Such an understanding has clearly important implications on strengthening critical network links.

Disruptions to the road network such as road repairs, or indeed with events of even greater magnitude and impact such as natural disasters, degrade its performance and cause the network to be less reliable by reducing its link capacities and thus increasing travel times. To assess the reliability of the network accurately and design a reliable transport network properly, Szeto et al. (2006) develop a risk-averse user equilibrium traffic assignment model with elastic demand by considering two sub-problems: the user problem and the demon problem. The user problem describes the non-cooperative and elastic behavior of the potential travel demand, whereas the demon problem describes the evil behavior in the sense of trying to cause maximum damage to the users. These two problems are combined as a nonlinear complementarity problem and solved by a modified branch and bound algorithm. A small numerical study is set up to illustrate the properties and the performance of the algorithm.

Sumalee et al. (2011) propose a multi-modal transport network assignment model considering uncertainties in both demand and supply sides of the network. These uncertainties are attributed to the adverse weather conditions with different degrees of impacts on different modes. The paper provides the derivations of mean and variance–covariance of the stochastic passenger flows and disutility terms involved in the route/mode choice model under the common-line framework. The risk-averse travelers are assumed to consider both the mean and variance of the random perceived travel time on each multi-modal path in their path choice decisions. The model also considers travelers' perception errors by using a Probit stochastic user equilibrium framework which is formulated as fixed point problem. A heuristic solution algorithm is proposed to solve the fixed point problem. Numerical examples are presented to illustrate the applications of the proposed model.

Table 3.1: The main details of the literature considering vehicle capacity

Authors	Method	Objectives	Constraints	Application	cap	s.cap	u.cap
Constantin & Florian (**1995)**	projected sub gradient algorithm	to minimize the passengers total expected travel and waiting time	fleet size	Ex: Stockholm, Winnipeg and Portland			*
Cominetti & Correa (**2001)**	hyperpath-Dijkstra method	minimize the expected transit time	the waiting times, the flow distribution	Small example	*	*	*
Lam et al. (**2002)**	stochastic user equilibrium transit assignment model	minimizes their perceived total travel time	invehicle capacity, line flow	Numerical example	*		
Tom & Mohan (**2003)**	genetic algorithms (GA)	minimize bus operating cost and passenger total travel time	operational constraints, network size, frequency feasibility, load factor	sample study on a medium-sized network	*		

Table 3.1: The main details of the literature considering capacity (Cont'd)

Authors	**Method**	**Objectives**	**Constraints**	**Application**	**cap**	**s.cap**	**u.cap**
Monyrath (**2006**)	Hooke and Jeeves algorithm	minimize the total travel cost of multi-modal network users	line frequency	Numerical example	*		
Leiva et al. (**2010)**	optimization programming language AMPL	minimizes wait time, in-vehicle travel time and operator cost	capacity, attractive lines constraint ability to transfer	Santiago, Chile	*****		*****
Yoo et al. (**2010**)	a gradient projection method	the upper problem is an operator's frequency design, the demand maximization and the lower level problem is a user's mode and route choice problem with variable demand	fleet size constraints, capacity, transfer delays between transit lines	Small example	*****		

Table 3.1: The main details of the literature considering capacity (Cont'd)

Authors	**Method**	**Objectives**	**Constraints**	**Application**	**cap**	**s.cap**	**u.cap**
Szeto & Wu (**2011)**	the route design and frequency setting problems with GA	to min.the number of transfers and total travel time of the passengers	the fleet size in-vehicle travel time, frequency requirement, intermediate stops	Ex: Tin Shui Wai, Hong Kong			*
Yu et al. (**2011)**	parallel genetic algorithm (PGA),	the maximization of service quality and the minimization of operational costs.	vehicle fleet size	Dalian City, China			*
Hamdouch et al. (**2011)**	method of successive averages (MSA)	to differentiate the discomfort level experienced by the sitting and standing passengers	First-Come-First-Serve (FCFS) principle, travel time, seat and standing capacities	Numerical example	*	*	

Table 3.1: The main details of the literature considering capacity (Cont'd)

Authors	**Method**	**Objectives**	**Constraints**	**Application**	**cap**	**s.cap**	**u.cap**
Gallo et al. (**2011)**	a new heuristic method and a meta-heuristic algorithm based on scatter search	external costs and private car costs	capacity, assignment, multimodal assignment constraint	small network and on a real-scale network	*		*
Hadas & Shnaiderman (**2012**)	frequency setting with the use of stochastic properties	to minimize the total cost with decision variables of either frequency or vehicle capacity	the fleet size, frequency requirement	Small example	*	*	

(Table 2.1 summarize the studies such as: cap = capacitated problem, s.cap = seat capacitated problem, u.cap = uncapacitated problem)

With the aim to examine the role of risk perception in use of private and public modes of transportation, Rundmo et al. (2011) carry out a mailed self-completion questionnaire survey among a representative sample of the Norwegian public aged from 18 to 65 years (n = 1864). Perceived control related to private modes of transportation, knowledge about safety and trust in authorities were found to be significantly different among respondents who often used private modes of transportation compared to those who most often used public modes. Additionally, no significant difference is found in the severity of consequences due to which transport modes that the respondents used most frequently. The study concludes that the role of consequence judgment for precautionary action and demand for risk reduction are misleading when generalized to decisions about transport mode use.

A state-of-the-art review of the transport network design problem (NDP) under uncertainty is given by Chen et al. (2011). In their paper, the authors also present a bi-objective-reliable NDP (BORNDP) model that explicitly optimizes the capacity reliability and travel time reliability under demand uncertainty. These performance measures are useful as they can describe the supply-side reliability and demand-side reliability of a road network. A simulation-based multi-objective genetic algorithm solution procedure, which consists of a traffic assignment algorithm, a genetic algorithm, a Pareto filter, and a Monte-Carlo simulation, is developed to solve the proposed BORNDP model. A numerical example based on the capacity enhancement problem is presented to demonstrate the tradeoff between capacity reliability and travel time reliability in the NDP.

In their recent study on transit systems, Ceder et al. (2013) claim that out-of-vehicle times are perceived as being more onerous than in-vehicle time by transit users when making transfers. To back their claims, authors conduct a user preference survey at to major terminals to determine the effects of uncertainty in out-of-vehicle times during transfers on the willingness to use transfer routes for transit users. They also aim determine the influence of out-of-vehicle facilities offered by public- transport operators on transit users' perception of trip attributes related to transfers. The survey data are modeled using cumulative prospect theory and fuzzy logic. The results show that for all trip attributes, except for comfort, transit users' exhibit risk-averse behavior. Moreover, findings indicate that transit users who are accustomed to better out-of-vehicle facilities have a lower tolerance for uncertainty in transfer waiting times and delay times.

3.2 Measuring Risks: Conditional Value-at-Risk (CVaR)

In the work influential work of Artzner et al. (cvar thesis 7) a clear definition of a *coherent risk measure* is given based on four axioms. Within a financial perspective, let X and Y denote portfolio returns, $\rho(X)$ and $\rho(Y)$ are the risk measures of these portfolios respectively, and c a constant. Then, a coherent risk measure must fulfill the following four axioms:

Positive Homogeneity	$\rho(cX) = c\rho(X)$,
Subadditivity	$\rho(X + Y) \leq \rho(X) + \rho(Y)$,
Transitional invariance	$\rho(X + c) = \rho(X) - c$,
Monotonicity	$\rho(X) \leq \rho(Y)$, if $X \leq Y$.

The first axiom means that the risk of a portfolio is proportional to its size. The second axiom ensures that the total portfolio risk is less than or equal to the sum of the risks of its components. This is why in fact diversification in a portfolio is a praised characteristic. The third axiom implies that adding some cash to the portfolio decreases its risk by the same amount. Finally, the last axiom implies that if the return of portfolio Y dominates the return of portfolio X, then the risk of portfolio Y cannot be lower than the risk of portfolio X. Some authors replace the first two axioms of coherence with condition that ρ be convex, or

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y) \tag{3.1}$$

for $0 \leq \lambda \leq 1$. However, convexity does not necessarily imply positive homogeneity. Thus a risk measure that is only convex, monotone and transitionally invariant has weaker properties and called *weak coherent*.

Nobel prize winner Markowitz was the first researcher to propose variance, the deviation from the mean of the return distribution, as a risk measure. However, when variance is used as a risk measure, gains and losses are equally penalized. Additionally, mean-variance decisions are usually not consistent with the expected utility approach with the exception that normally distributed or a quadratic utility function is used. Finally, the variance does not account for fat tails of the underlying distribution and thus it is inappropriate to describe the risk of low probability events.

To respond to the need to aggregate the various sources of risks of a portfolio, the concept of value-at-risk (VaR) was introduced in 1994. In simple terms, it is defined as the maximum loss with probability α over a certain time horizon. More formally, let Z be a random variable of loss type with a cumulative distribution function denoted as F_Z. Then,

$$\mathrm{VaR}_\alpha(Z) = \inf\{\mu \ : \ F_Z(\mu) \geq \alpha\} \tag{3.2}$$

at confidence level $\alpha \in [0, 1)$. Despite its wide acceptance in the literature and practice, VaR does not fulfill the axioms of coherence. The most significant violation is related with subadditivity, in other words the risk of a portfolio measured with VaR may be larger than the sum of risks of its components. Another critique on VaR is due to its non-convex characteristic. This limits its use as a risk measure in optimal portfolio selection for investment. Indeed, it has been shown that using VaR may incur larger losses in the most adverse states compared to risk neural case (cvar thesis 5). Other criticisms on VaR include its inability to measure losses exceeding VaR, its unsuitability for ranking as it has many local extremes and its conflicting results at different confidence levels.

In general, variance and VaR are not coherent and lead to meaningless results and also do not allow to measure the degree of co-dependence (positive or negative) between the random variables in the case of non-elliptic (but possibly symmetric) joint probability distributions. To investigate tail events, Embrechts et al. (1997) [Szegö] introduced the concept of α-expected shortfall or α-tail mean. Later on, Uryasev (2000) [Szegö] presented a similar measure, namely conditional value-at-risk (CVaR). It is defined as

$$\mathrm{CVaR}_\alpha(Z) \ = \inf\left\{\eta + \ \frac{1}{1-\alpha}\mathbf{E}([Z-\eta]_+) \ : \ \eta \in \mathbb{R}\right\} \tag{3.3}$$

Where $[z]_+ = \max(x, 0)$ denotes the positive part of a number $x \in \mathbb{R}$. CVaR is also known in the literature as average value-at-risk and tail value-at-risk, due to the following expression

$$\mathrm{CVaR}_\alpha(Z) \ = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_\gamma(Z)d\gamma. \tag{3.4}$$

Suppose that Z denotes the random total travel cost. Then, $\mathrm{VaR}_\alpha(Z)$ is the α-quantile (a high quantile) of the distribution of the total travel cost, which provides an upper bound on the cost that is exceeded only with a small probability of $1-\alpha$. On the other hand, $\mathrm{CVaR}_\alpha(Z)$ is a measure of severity of the cost if it is larger than $\mathrm{VaR}_\alpha(Z)$. It is easy to see from (3.4) that $\mathrm{CVaR}_\alpha(Z)$ is related to the expectation of the cost exceeding the threshold $\mathrm{VaR}_\alpha(Z)$. This is also depicted in Fig 3.1.

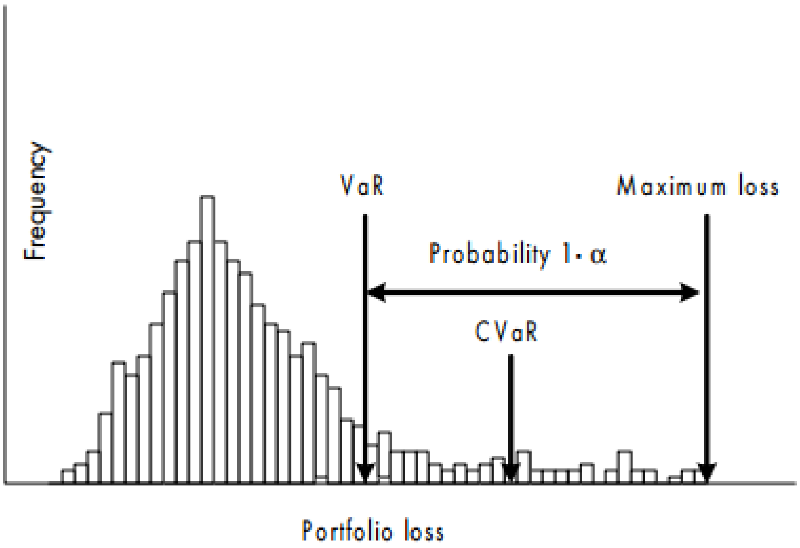

Figure 3.1: VaR and CVaR

Let Z be a random variable with (not necessarily distinct) realizations $z_1, \dots, z_n$ and corresponding probabilities $p_1, \dots, p_n$. Then, for a given confidence level $\alpha \in [0, 1)$ the optimum value of the following optimization problem is equal to $\mathrm{CVaR}_\alpha(Z)$

$$\min \quad \mu + \frac{1}{1-\alpha}\sum_{i=1}^{n} p_i w_i \tag{3.5a}$$

$$\text{s.t.} \quad w_i \geq z_i - \mu \qquad i = 1, \dots, n, \tag{3.5b}$$

$$w_i \geq 0 \qquad i = 1, \dots, n. \tag{3.5c}$$

There are very few studies in the available literature that make use of CVaR risk measure in the context of transportation. Chen and Zhou (2010) propose the α-reliable mean-excess traffic equilibrium model (METT) that explicitly considers both reliability and unreliability aspects of travel time variability in the route choice decision process. In contrast to the travel time budget (TTB) models that consider only the reliability aspect defined by TTB, METT hypothesizes that travelers are willing to minimize their mean-excess travel times defined as the conditional expectation of travel times beyond the TTB. Although authors do not use CVaR risk measure in their model, they claim that METT is consistent with CVaR. The model is formulated as a variational inequality problem and solved by a route-based traffic assignment algorithm via the self-adaptive alternating direction method. Illustrative examples are presented to demonstrate the characteristics of the model as well as its differences compared to TTB models.

Seyedshohadaie et al. (2010) present a method for determining optimal risk-based maintenance and rehabilitation policies for transportation infrastructure. Under a predefined level of risk, the proposed policies guarantee a certain performance level across the network. The long-term model is formulated in the Markov Decision Process framework with risk-averse actions and transitional probabilities describing the uncertainty in the deterioration process. CVaR is used as the measure of risk and policies are modeled assuming no budget restriction. Two linear programming models are formulated to generate network-level polices with different objectives to address the short-term resource allocation problem. The proposed methodology is general and can be used with any performance indicator. Numerical studies are based on pavement roughness and an analytical expression for computing CVaR is derived.

Azad and Davoudpour (2013) give a stochastic supply chain network design model in which three levels compromised of suppliers, distribution centers and customers seek to determine their optimal plans. CVaR is used as a risk measure to minimize large realizations of the annual cost. The problem is formulated as a convex mixed integer program and a two-phase (construction and improvement) heuristic method is developed to solve the problem. An initial solution is built randomly in the construction phase. The initial solution is then improved iteratively in improvement phase by using a hybrid algorithm combining Tabu search and simulated annealing methods. The authors validate their model by comparing CVaR measure with mean-variance measure.

Sorokin et al. (2013) consider a formulation for the fixed charge network flow problem subject to multiple uncertain arc failures. Their aim is to provide a robust optimal flow assignment in the sense of restricting potential losses using CVaR. The authors show that a heuristic algorithm referred to as Adaptive Dynamic Cost Updating Procedure can be extended to the considered problem under uncertainty and produce high-quality heuristic solutions for large problem instances.

The last study that we could access is due to Toumazis and Kwon (2013). The authors propose a new method for mitigating risk in routing hazardous materials (hazmat), based on CVaR measure on time-dependent vehicular networks. Specifically, they study the problem in which accident probabilities and accident consequences are time-dependent; that is, the probability of an accident and the resulting consequences depend on the shipment's entrance time in the arc mainly due to traffic condition. They also provide a numerical method to determine an optimal departure time and an optimal route for a given origin–destination pair. The authors show that the CVaR models are flexible and suitable for hazmat transportation.

It can be remarked that as a risk measure, CVaR is very recently used in transportation problems and the obtained results are very encouraging. The traditional objective to transit line frequency setting is the minimization of total travel time given forecasted mean values. This approach occasionally leads to excessive delays for the passengers due to the changing conditions. In this study, the random network parameters are described by using a finite set of scenarios and a risk-averse mathematical model is developed using CVaR. As CVaR has never been used in planning transit system, this study means an important step for forthcoming research.

3.3 Bi-Level Programming

Many real-world problems involving a hierarchical relationship between decision levels can be modeled as multilevel programs. This type of problems are encountered in many fields such as management (facility location, credit allocation, hazardous materials, energy policy, environmental regulation,), economic planning (oil production, social and agricultural policies, electric power pricing), engineering (optimal design, structures and shape), environmental sciences, chemistry, optimal control, etc. A small set of example studies is provided is here including revenue management (Cote et al., 2003), congestion management (Hearn and Ramana, 1998), origin-destination matrix estimation (Florian and Chen, 1995), management of hazardous materials (Kara and Verter, 2004), network design problems (Constantin and Florian, 1995), energy sector (Hobbs and Nelson (1992).

Let us consider the toll-setting problem where the objective is to maximize the revenue obtained from the toll-set of a transportation network. In general network users' objective is to minimize their travel cost (or in the simplest case, their travel time). Therefore at optimality, toll levels should not be set too high to not direct users avoid tolled arcs but still generating large revenues. Once the tolls are set, users react and select their route so as to minimize their total travel cost. This hierarchical relationship between the two independent decision makers having probably diverging point of views is the significant characteristic of this problem. In economics theory, it is related with the Stackelberg, or leader-follower game (Colson et al., 2007).

Let us denote by $\mathcal{A}$ the set of links of the network and by $\bar{\mathcal{A}}$ the subset of toll links. Then, the mathematical program related with the network manager is

$$\max_{T,x} \sum_{a \in \overline{\mathcal{A}}} T_a x_a \tag{3.6a}$$

$$\text{s.t.} \quad l_a \leq T_a \leq u_a \qquad\qquad a \in \bar{\mathcal{A}}. \tag{3.6b}$$

where T_a denotes the toll level, x_a the flow, l_a and u_a lower and upper bounds respectively on toll levels for link a. Let us assume that network users adopt a selfish behavior. Then, the system reaches to an equilibrium where all users are assigned to paths of minimum cost with respect to the current congestion levels. In a congestion-free environment such user equilibrium coincides with a flow assignment that minimizes total system cost. It follows that the path-flow vector $\mathbf{f}$ together with the link-flow vector $\mathbf{x}$ is solution of the linear program:

$$\min_{f,x} \sum_{a\in\mathcal{A}} c_a x_a + \sum_{a\in\bar{\mathcal{A}}} T_a x_a \tag{3.7a}$$

$$\text{s.t.} \quad \sum_{p\in P_{rs}} f_p^{rs} = d_{rs} \qquad (r,s)\in\mathcal{OD}, \tag{3.7b}$$

$$x_a = \sum_{(r,s)\in\mathcal{OD}} \sum_{p\in\mathcal{P}_{rs}} \delta_{a,p}^{rs} f_p^{rs} \qquad a\in\mathcal{A}, \tag{3.7c}$$

$$f_p^{rs} \geq 0 \qquad p\in\mathcal{P}_{rs},\ (r,s)\in\mathcal{OD}. \tag{3.7d}$$

In the above model, the objective (3.7*a*) sum up tolls $T_a\,(a\in\bar{\mathcal{A}})$ and the costs c_a for each link. Constraint (3.7*b*) represents demand satisfaction for an origin and destination pair $(r,s)\in\mathcal{OD}$ where $\mathcal{OD}$ is the set of all origin-destination pairs. Constraint (3.7*c*) links path flows $f_p^{r,s}$ and link flows x_a with

$$\delta_{a,p}^{rs} = \begin{cases} 1 & \text{if path } p\in\mathcal{P}_{rs} \text{ includes link } a, \\ 0 & \text{otherwise.} \end{cases}$$

Mathematical programming models (3.6) and (3.7) are connected due to the use of common variables $T_a\,(a\in\bar{\mathcal{A}})$ and flows $x_a\,(a\in\mathcal{A})$. Moreover, the profit of the network manager given in (3.6*a*) cannot be calculated until the network flows are known. These flows are the solution of a mathematical program parameterized in the toll vector $\mathbf{T}$. Hence, the bilevel formulation in (3.8) is built.

$$\min_{T,f^*,x^*} \quad \sum_{a\in\bar{\mathcal{A}}} T_a x_a^* \tag{3.8a}$$

$$\text{s.t.} \quad l_a \leq T_a \leq u_a \qquad a \in \bar{\mathcal{A}}, \tag{3.8b}$$

$$(f^*, x^*) \in \arg\min_{f,x} \quad \sum_{a\in\mathcal{A}} c_a x_a + \sum_{a\in\overline{\mathcal{A}}} T_a x_a \tag{3.8c}$$

$$\text{s.t.} \quad \sum_{p\in P_{rs}} f_p^{rs} = d_{rs} \qquad (r,s) \in \mathcal{OD}, \tag{3.8d}$$

$$x_a = \sum_{(r,s)\in\mathcal{OD}} \sum_{p\in\mathcal{P}_{rs}} \delta_{a,p}^{rs} f_p^{rs} \qquad a \in \mathcal{A}, \tag{3.8e}$$

$$f_p^{rs} \geq 0 \qquad p \in \mathcal{P}_{rs},\ (r,s) \in \mathcal{OD}. \tag{3.8f}$$

The hierarchy is a consequence of the fact that the mathematical program related to behavior of the network users is part of the manager's constraints. This is the main characteristic of bilevel programs: they include two mathematical programs within a single instance, one of these problems being part of the constraints of the other one. Within this hierarchy, the program (3.6) is called the *upper-level problem* while (3.7) corresponds to the *lower-level problem*.

The general formulation of a bi-level programming problem is (Colson et al., 2007),

$$\min_{\mathbf{x}\in X,\mathbf{y}} \quad F(\mathbf{x},\mathbf{y}) \tag{3.9a}$$

$$\text{s.t.} \quad \mathbf{G}(\mathbf{x},\mathbf{y}) \leq 0 \tag{3.9b}$$

$$\min_{\mathbf{y}} \quad f(\mathbf{x},\mathbf{y}) \tag{3.9c}$$

$$\text{s.t.} \quad \mathbf{g}(\mathbf{x},\mathbf{y}) \leq 0 \tag{3.9d}$$

where $\mathbf{x} \in \mathbb{R}^{n_1}$, $F: \mathbb{R}^{n_1} \text{ x } \mathbb{R}^{n_2} \rightarrow \mathbb{R}$ and $\mathbf{G}: \mathbb{R}^{n_1} \text{ x } \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{m_1}$ are the upper level variables, objective function and constraints respectively. Similarly, $\mathbf{y} \in \mathbb{R}^{n_2}$, $f: \mathbb{R}^{n_1} \text{ x } \mathbb{R}^{n_2} \rightarrow \mathbb{R}$ and $\mathbf{g}: \mathbb{R}^{n_1} \text{ x } \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{m_2}$ are the lower level variables, objective function and constraints respectively. Upper-level constraints involve variables from both levels and play a very specific role. Indeed, they must be enforced indirectly, as they do not bind the lower-level decision-maker.

3.4 Model Formulation

Analogous to Florian (2003), let's assume that the network $G = (\mathcal{N}, \mathcal{A})$ is composed of four types of arcs: wait arcs (no travel time), in-vehicle (no waiting), alighting (no travel and waiting time) and walk arcs (travel time, no waiting). It is assumed that the underlying network is strongly connected. The segment of a transit line is an arc that is served by a vehicle at given intervals. The transit traveler waits for the link to be served by a vehicle. At each node that is on the itinerary of a transit line, the distribution of the interarrival times of the vehicles is known for each line which serves the node. As a consequence, on can compute the combined time for the arrival of the first vehicle, for any subset of lines incident at a node, as well as the probability that each line arrives first.

a. Sets / Indices:

$a \in \mathcal{A}$	Set of links
$i \in \mathcal{N}$	Set of nodes
$d \in \mathcal{D}$	Set of destinations
$l \in \mathcal{L}$	Set of lines
$s \in \mathcal{S}$	Set of scenarios

b. Parameters:

M	Available fleet size
$\underline{f_l}, \overline{f_l}$	Minimum and maximum service frequency respectively for line l
t_l	Travel time for line l
α	Confidence level
p_s	Realization probability for scenario s
g_{id}^s	Travel demand from origin i to destination d for a scenario s
δ_{al}	Indicator with value 1 if link a is on line l, and 0 otherwise
t_a^s	Travel time on arc a for scenario s
t_l^s	Travel time for line l and scenario s $(= \sum_a \delta_{al} t_a^s)$
r^s	Realization or total cost (travel + waiting) for scenario s

c. Variables:

f_l	Frequency of line l
v_{ad}^s	Flow on link a heading towards destination d for scenario s
w_{id}^s	Total waiting time of passengers heading towards destination d at node i for scenario s
η	VaR value at confidence interval α
c^s	Auxiliary variable such that $c^s = [r^s - \eta]_+$ for scenario s

3.4.1 Lower Level Problem

For simplicity, we assume for now that there is only one destination q. We denote the set arcs that will be included in a solution as $\bar{\mathcal{A}} \subseteq \mathcal{A}$, and accordingly the solution for a single destination q is denoted as a acyclic subgraph $b_q = (\mathcal{N}, \bar{\mathcal{A}})$. The travel demand from node $i \in \mathcal{N}$ to destination q is denoted as g_i. $\bar{\mathcal{A}}$, A traveler at node is $i \in \mathcal{N}$ is assumed to board the first vehicle operating on any of the arcs in $\bar{\mathcal{A}}_i^+$ where $\bar{\mathcal{A}}_i^+ \subseteq \bar{\mathcal{A}}$ corresponds to the forward star of node i or also referred as the *set of attractive lines*.

Let $W(\bar{\mathcal{A}}_i^+)$ be the expected waiting time for the arrival of the first vehicle serving any of the links $a \in \bar{\mathcal{A}}_i^+$, which is denoted as the *combined waiting time* of links $a \in \bar{\mathcal{A}}_i^+$. Let $P_a(\bar{\mathcal{A}}_i^+)$ be the probability that link a is the first line to be served among the links $\bar{\mathcal{A}}_i^+$. If an exponential distribution of interarrival time is supposed and f_a is the frequency of line a, then

$$W(\bar{\mathcal{A}}_i^+) = \frac{1}{\displaystyle\sum_{a \in \bar{\mathcal{A}}_i^+} f_a} \tag{3.10}$$

and

$$P(\bar{\mathcal{A}}_i^+) = \frac{f_a}{\displaystyle\sum_{a \in \bar{\mathcal{A}}_i^+} f_a}, \qquad a \in \bar{\mathcal{A}}_i^+ \tag{3.11}$$

Since $\bar{\mathcal{A}}$ is not known a priori the single destination model is formulated by using binary variables x_a

$$x_a = \begin{cases} 0 & \text{if } a \notin \bar{\mathcal{A}}, \\ 1 & \text{if } a \in \bar{\mathcal{A}}. \end{cases} \tag{3.12}$$

The optimization model may be stated as follows.

$$\min \quad \sum_{a\in\mathcal{A}} t_a v_a + \sum_{i\in N} \frac{V_i}{\sum\limits_{a\in\mathcal{A}_i^+} f_a x_a} \tag{3.13a}$$

$$\text{s.t.} \quad v_a = \frac{x_a f_a}{\sum\limits_{a'\in\mathcal{A}_i^+} f_{a'} x_{a'}}, \qquad a\in\mathcal{A}_i^+,\ i\in\mathcal{N} \tag{3.13b}$$

$$V_i = \sum_{a\in\mathcal{A}_i^+} v_a + g_i, \qquad i\in\mathcal{N} \tag{3.13c}$$

$$V_i \geq 0, \qquad i\in\mathcal{N} \tag{3.13d}$$

$$x_a = \{0,1\}, \qquad a\in\mathcal{A}, \tag{3.13e}$$

where t_a is the travel cost (time) on link a and V_i is the total volume at node i. The problem in (3.13) is a mixed integer nonlinear optimization problem. Hopefully, it can be reduced to a simpler linear programming problem as constraint (3.14*d*) may be replaced by the nonnegavity constraints of the link volumes $v_a \geq 0 \quad a\in\mathcal{A}$ since $V_i = \sum_{a\in\mathcal{A}_i^+} v_a \quad i\in N$. Therefore, by setting $w_i = V_i / \left(\sum_{a\in\mathcal{A}_i^+} f_a x_a\right)$ for all $i\in\mathcal{N}$, where the new variable w_i corresponds to the total waiting time of all trips at node i, one can obtain the following model

$$\min \quad \sum_{a\in\mathcal{A}} t_a v_a + \sum_{i\in N} w_i \tag{3.14a}$$

$$\text{s.t.} \quad v_a = x_a f_a w_i, \qquad a\in\mathcal{A}_i^+,\ i\in\mathcal{N} \tag{3.14b}$$

$$\sum_{a\in\mathcal{A}_i^+} v_a - \sum_{a\in\mathcal{A}_i^-} v_a = g_i, \qquad i\in\mathcal{N} \tag{3.14c}$$

$$v_a \geq 0, \qquad a\in\mathcal{A}. \tag{3.14d}$$

The objective function (3.14*a*) is linear. The binary variables are only used in the nonlinear constraint (3.14*b*) which can be relaxed by

$$v_a \leq f_a w_i, \qquad a \in \mathcal{A}_i^+, i \in \mathcal{N}. \tag{3.15}$$

With this final modification, the model described with (3.14*a*), (3.15), (3.14*c*) and (3.14*d*) becomes a linear programming (Florian, 2003). This last model can be easily extended to the case where there are multiple destinations $d \in \mathcal{D}$ and multiple line on a link $a \in \mathcal{A}$.

$$\min \quad \sum_{d \in \mathcal{D}} \Big(\sum_{a \in \mathcal{A}} t_a v_{ad} + \sum_{i \in N} w_{id} \Big) \tag{3.16a}$$

$$\text{s.t.} \quad v_{ad} \leq \Big(\sum_{l \in \mathcal{L}} \delta_{al} f_l \Big) w_{id}, \qquad a \in \mathcal{A}_i^+,\ i \in \mathcal{N},\ d \in \mathcal{D} \tag{3.16b}$$

$$\sum_{a \in \mathcal{A}_i^+} v_{ad} - \sum_{a \in \mathcal{A}_i^-} v_{ad} = g_{id}, \qquad i \in \mathcal{N},\ d \in \mathcal{D} \tag{3.16c}$$

$$v_{ad} \geq 0, \qquad a \in \mathcal{A},\ d \in \mathcal{D}. \tag{3.16d}$$

In model (3.16), v_{ad} is the flow on link a heading towards destination node d, w_{id} is the total waiting time of passengers at node i heading towards destination node d and g_{id} is travel demand from origin node i to destination node d.

In this study, we characterize the inherent uncertainty by a set of scenarios and generalize the model in (3.16). Two types of uncertainty are considered: time and demand. When buses do not operate on dedicated lanes and share the transportation infrastructure with other vehicles, the congestion affects their travel times. Even operating on dedicated lanes, adverse conditions such as road constructions or bad weather may be the sources of delays for buses. Travel demand from one location to another is also uncertain by nature. In-day or day-to-day changes in demand are common. Here, we focus only on the peak hour of the workdays.

$$\min \quad \sum_{s\in\mathcal{S}} p_s \left[\sum_{d\in\mathcal{D}}\Big(\sum_{a\in\mathcal{A}} t_a^s v_{ad}^s + \sum_{i\in\mathcal{N}} w_{id}^s\Big)\right] \tag{3.17a}$$

$$\text{s.t.} \quad v_{ad}^s \leq \Big(\sum_{l\in\mathcal{L}} \delta_{al} f_l\Big) w_{id}^s, \qquad a\in\mathcal{A}_i^+,\ i\in\mathcal{N},\ d\in\mathcal{D},\ s\in\mathcal{S} \tag{3.17b}$$

$$\sum_{a\in\mathcal{A}_i^+} v_{ad}^s - \sum_{a\in\mathcal{A}_i^-} v_{ad}^s = g_{id}^s, \qquad i\in\mathcal{N},\ d\in\mathcal{D},\ s\in\mathcal{S} \tag{3.17c}$$

$$v_{ad}^s \geq 0, \qquad a\in\mathcal{A},\ d\in\mathcal{D},\ s\in\mathcal{S}. \tag{3.17d}$$

The model in (3.17) has the objective to minimize the expected cost, or the total time that passengers travel in the vehicles and wait at the stations. The constraint (3.17*b*) is based on the assumption that all passengers can get on the first arriving vehicle, in other words vehicle capacity is not limited. In that case, this constraint builds the relationship between the flow on a link emanating for a node i and the waiting time of passengers in that node based on the total frequency on that link. The constraint (3.17*c*) corresponds to the traditional flow balance constraint for the network flow models, where the sum of flows leaving node i must be equal to the sum of flows arriving to node i from other nodes and the passengers existing at node i. The last constraint (3.17*d*) ensures the non-negativity of link flows, and also the non-negativity of the waiting times indirectly.

3.4.2 Upper Level Problem

In the upper level problem, we use a risk measure to take into consideration the effect of the stochastic nature of the system and model the network reliability. In particular, we consider CVaR as an asymmetric risk measure on the function of the random travel times and demands. We aim to minimize CVaR value of the system wide passengers' total travel cost (time) in order to control its realizations that are above certain threshold values. Trying to find a policy that avoids large realizations of the specified random performance measure is consistent with the objectives of the risk averse transportation planning manager. We assume that frequency setting decisions do not vary according to the stochastic nature of the network and the users have no information except the probability distributions about the network conditions while choosing their routes. In other words, the users have no information about the network realization and observe the conditions after making their route choices, and therefore, they cannot vary their decisions according to the network realization. Let us denote the realization of scenario s as

$$r^s = \sum_{d \in \mathcal{D}} \Big(\sum_{a \in \mathcal{A}} t_a^s v_{ad}^s + \sum_{i \in \mathcal{N}} w_{id}^s \Big). \tag{3.18}$$

Using the definition given in (3.5), we arrive at the upper level model minimizing CVaR of the total system cost at a specified confidence level α:

$$\min \quad \eta + \frac{1}{1-\alpha} \sum_{s \in \mathcal{S}} p_s \, c^s \tag{3.19a}$$

$$\text{s.t.} \quad c^s \geq r^s - \eta, \qquad s \in \mathcal{S} \tag{3.19b}$$

$$c^s \geq 0, \qquad s \in \mathcal{S} \tag{3.19c}$$

$$\sum_{l \in \mathcal{L}} t_l^s f_l \leq M, \qquad s \in \mathcal{S} \tag{3.19d}$$

$$\underline{f}_l \leq f_l \leq \overline{f}_l, \qquad l \in \mathcal{L}. \tag{3.19e}$$

In this formulation, c^s, $s \in \mathcal{S}$, variables are introduced to calculate the realization of random variable $[r^s - \eta]_+$ under each scenario. By constraints (3.19*b*) and (3.19*c*), and the nature of the objective function it is guaranteed that $c^s = [r^s - \eta]_+$ for all $s \in \mathcal{S}$. Then, by definition of CVaR given in (3.3) the optimal value of the objective function (3.19*a*) is equal to CVaR(r), as desired. Constraints (3.19*d*) are for avoiding unrealistic frequency assignments exceeding available fleet size. The last constraints

(3.19*e*) are so called service constraints as they impose minimum and maximum service levels for line frequencies.

In our model, the variables f_l, $l \in \mathcal{L}$, are upper level variables whereas v_{ad}^s, $a \in \mathcal{A}$, $d \in \mathcal{D}$, $s \in \mathcal{S}$, and w_{id}^s, $i \in \mathcal{N}$, $d \in \mathcal{D}$, $s \in \mathcal{S}$ are the lower level variables. The system wide cost for the network manager cannot be calculated until $\mathbf{v}$ and $\mathbf{w}$ are known. These values are the solution of the mathematical program in (3.17) parameterized in the frequency vector $\mathbf{f}$.

4 PROPOSED SOLUTION METHOD

4.1 Solving the Lower Level Model

For simplicity, let us focus on the linear programming model described with (3.14*a*), (3.15), (3.14*c*) and (3.14*d*). The dual of this program is

$$\max \quad \sum_{i\in\mathcal{N}} g_i u_i \tag{4.1a}$$

$$\text{s.t.} \quad u_j + t_a + \mu_a \geq u_i, \qquad a \in \mathcal{A} \tag{4.1b}$$

$$\sum_{a\in\mathcal{A}_i^+} f_a\, \mu_a = 1, \qquad i \in \mathcal{N} \tag{4.1c}$$

$$\mu_a \geq 0, \qquad a \in \mathcal{A} \tag{4.1d}$$

where u_i and u_j are dual variables related with constraint (3.14*c*), and μ_a are the dual variables related with the constraint (3.14*b*). Let $(v^*,\ w^*)$ denote the optimum solution of the primal problem in (3.14) and $(u^*,\ \mu^*)$ denote the optimum solution of the dual problem in (4.1). We can then write the weak complementary slackness conditions such as

$$(v_a^* - f_a w_i^*)\, \mu_a^* = 0, \quad a \in \mathcal{A}_i^+,\ i \in \mathcal{N} \tag{4.2}$$

and

$$(u_a^* + t_a + \mu_a^* - u_a^*)\, v_a^* = 0, \quad a \in \mathcal{A}. \tag{4.3}$$

Given the primal and dual formulations, it can be observed that the transit route choice problem (TRCP) has a close similarity to the shortest path route choice problem. Hence, a solution algorithm which solves TRCP can be built in resemblance to the label setting algorithm for computing shortest paths. In the first stage of this algorithm, the set of arcs which carry flow, $\bar{\mathcal{A}}$, and the expected travel times u_i^* from each node $i \in \mathcal{N}$ to the destination node is computed. Then the demand from all origins to the destination is assigned to the arcs $a \in \bar{\mathcal{A}}$ in the second stage. The algorithm that solves TRCP with a single destination node q is given in Algorithm 1 below (Florian, 2003). As the lower level problem given in (3.17) is a linear programming problem that can be decomposed in destinations and scenarios, we run $|\mathcal{D}| \times |\mathcal{S}|$ times Algorithm 1 to solve a single instance of the problem. Note that we use the convention $0.\infty = 1$.

4.2 Solving the Bilevel Problem

It was not until the early nineteen eighties that the value of the bilevel programs in modelling hierarchical decision processes and engineering design problems prompted researchers to pay close attention. As bilevel programming problems are fundamentally difficult, most algorithmic research is focused at first on the simplest cases of bilevel programs, that is problems having nice properties such as linear, quadratic or convex objective and/or constraint functions. In particular, the most studied instance of bilevel programming problems is the linear bilevel programs in which all functions are linear. Several surveys such as those by Hsu and Wen (1989), Wen and Hsu (1991) and Ben-Ayed (1993) are dedicated to this class.

Algorithm 1: Algorithm to solve TRCP in (4.1)

Step 1: $u_i = \infty$ for all $i \in \mathcal{N}/\{q\};\ u_q = 0;\ \bar{f}_i = 0$ for all $i \in \mathcal{N};\ \mathcal{B} = \mathcal{A};\ \bar{\mathcal{A}} = \emptyset$;

Step 2: **if** $\mathcal{B} = \emptyset$ **then**

go to Step 3;

otherwise

find $a = (i, j)$ such that $(u_j + t_a)$ is the smallest value of $\mathcal{B}$;

$\mathcal{B} = \mathcal{B}/\{a\}$;

if $u_i \geq u_j + t_a$ then

$u_i = \left(\bar{f}_i u_i + f_a(u_i + t_a)\right) / \left(\bar{f}_i + f_a\right)$;

$\bar{f}_i = \bar{f}_i + f_a$;

$\bar{\mathcal{A}} = \bar{\mathcal{A}} \cup \{a\}$;

go to step 2;

endif

endif

Step 3: $V_i = g_i^q$ for all $i \in \mathcal{N}$;

for each $a \in \bar{\mathcal{A}}$ in decreasing order of $(u_j + t_a)$ **do**

$v_a^q = (f_a/\bar{f}_i)V_i$;

$V_j = V_j + v_a^q$;

endfor

for each $a \in \mathcal{A}/\bar{\mathcal{A}}$ **do**

$v_a^q = 0$;

endfor

$w_i^q = V_i/\bar{f}_i$ for all $i \in \mathcal{N}$;

More complex bilevel programs such as the one due to Vicente et al. (1996) having discrete variables are studied over the years. These studies are included to more general surveys as for example to Anandalingam and Friesz (1992). Dempe (2003) wrote an annotated bibliography on both nonlinear bilevel programming problems and mathematical programs with equilibrium constraints. The combinatorial nature of bilevel programming is reviewed in Marcotte and Savard (2005). The most recent study on bilevel programs is due to Colson et al. (2007). We refer the interested readers to this work for an overview on the solution methods for bilevel programs.

In recent years, there has been interest in solving the computationally challenging bilevel transportation problems with metaheuristic algorithms. Among the metaheuristics, genetic algorithms (GAs) have attracted a lot of attention due to their desirable properties such as the ability to explore large portions of the feasible region, ease of parallelization, and a demonstrated effectiveness for a variety of optimization problems. GAs are computationally simple, easy to implement, and powerful search procedures (Yin, 2000). In addition, they are not limited by restrictive assumptions about the search space such as continuity and the existence of derivatives. We are also motivated by the fact that GAs have been successfully applied to various transportation problems including network design (see, e.g., Drezner and Salhi, 2002; Drezner and Wesolowsky, 2003; Karoonsoontawong and Waller, 2006; Dimitriou and Stathopoulos, 2008; Mathew and Shrama, 2009; Xu et al., 2009), traffic signal control (see, e.g., Ceylan and Bell, 2004), network design and frequency setting (Aggarwal and Mathew, 2004), network design and toll pricing (Dimitriou et al., 2008), and pricing (see, e.g., Yin, 2000; Shepherd and Sumalee, 2004; Zhang and Yang, 2004; Sumalee et al., 2005; Gardner et al., 2010).

GA is a search heuristic from the class of evolutionary algorithms that simulates the process of natural evolution such as inheritance, mutation, selection, and crossover for finding the solutions of optimization problems. In our implementations, we use the GA provided in the MATLAB R2012a Toolbox. MATLAB provides a rich and convenient programming environment for rapid algorithm design, and its parallel programming environment is especially crucial for computationally efficient implementations of parallelizable GAs. To speed up some tedious calculations related to the objective function calls, we also coded routines in C++ and interfaced with MATLAB through MEX facility.

We summarize the general framework of the solution methodology in Algorithm 2. The initial population is composed of frequency vectors whose components are randomly generated from uniform distributions on the interval $[\underline{\mathbf{f}}, \overline{\mathbf{f}}]$. If any of the frequency vector $\mathbf{f}$ is not feasible with respect to constraint (3.19*d*) or $\beta = \max_s\{\mathbf{t}^s\,\mathbf{f}\} > M$, it is restored to feasibility such that $\mathbf{f} = (M/\beta)\mathbf{f}$. For each frequency vector $\mathbf{f} \in \mathbb{R}^{|\mathcal{L}|}$, TRCP is solved $|\mathcal{D}| \times |\mathcal{S}|$ times for each destination and scenario with Algorithm 1 to obtain the optimal link flows and passengers' waiting times. Then, the optimal link flows and waiting times associated with each frequency vector are used to obtain the realization of the random total cost measure of interest (3.18) under each scenario. These realizations are used to calculate the CVaR of the performance measure. The CVaR associated with each frequency vector is in fact the raw fitness value of an individual. @fitscalingtop and @selectionroulette options of MATLAB are chosen to scale the raw fitness values and select the parents for reproduction, respectively. Top scaling scales the top individuals equally while roulette selection chooses parents by simulating a roulette wheel, in which the area of the section of the wheel corresponding to an individual is proportional to the individual's fitness. The selection algorithm generates a random number to select one of the sections; it basically uses a multinomial scheme with the success probabilities equal to the corresponding areas. Once the parents are selected for reproduction, crossover and mutation operators are used to form the next generation. In our study, the standard @crossoverintermediate option is chosen as a crossover operator to create children by taking a weighted average of the parents. Meanwhile, a special mutation operator is implemented; it alters some random number of components of a frequency vector by generating the new value of a selected f_l component from the uniform distribution on the interval $\left[\, \underline{f}_l,\, \overline{f}_l \,\right]$. If any child is not feasible with respect to constraint (3.19*d*) then it is restored to feasibility. The developed algorithm is elitist in the sense that some fraction of the current generation which corresponds to the individuals that has the best fitness values are guaranteed to survive to the next generation. The algorithm iterates until a predetermined number of generations is reached.

Algorithm 2: Algorithm to solve risk-averse transit frequency setting problem

(*initializations*) generate the initial population of feasible frequency vectors $\mathbf{f} \in \mathbb{R}^{|\mathcal{L}|}$ by sampling from the uniform distributions on the intervals $[\underline{\mathbf{f}}, \overline{\mathbf{f}}]$;

repeat

    for each individual of the current population **do**

        use Algorithm 1 to find the optimum solution of TRCP;

        evaluate the realization of each scenario with (3.18);

        calculate the corresponding CVaR;

    endfor

    apply selection, crossover and mutation to generate the new population;

until maximum number of generations is reached

output the best frequency vector $\mathbf{f}^*$;

5 NUMERICAL ANAYSIS

A small transit network to show the advantages of the model is compared with the expectation model in our study. We compare the CVaR value of the expectation model optimum solution with the CVaR value of our model and report the relative improvement. Also, we study the effects of parameters such as time, demand and fleet size on the model effectiveness. Two methods are used when scenarios are generated: first, we change the travel time on each link for which a randomly selected node is head or tail; second, we change the demand of some randomly selected origin-destination pairs.

In the following section we give some details on generating the problem instances. Then, in Section 5.2 we provide results to demonstrate the computational effectiveness of the proposed CVaR model. Also, we present numerical results of our algorithm to analyze how the optimal solutions change related with the input parameters. Finally, we discuss the computational study to compare the optimum solution of the CVaR value of the expectation model and the CVaR value of our model.

5.1 Generation of problem instances

We deal with the effects of demand on each node $i, i \in \mathcal{N}$ and $d, d \in \mathcal{D}$ and also we focus on time on each arc $a, a \in \mathcal{A}$ in the transit frequency setting problem to see the effects on the fleet size in the network M.

After several trials on the model the most significant solutions are found and considered for each scenario, respectively;

- ✓ demand on each bus stop,
- ✓ time on each arc, and
- ✓ fleet size values.

Here, we use three nodes to show the results of our model. There is one destination point and two line that goes to the destination point. For each demand points in the network problem instances are generated randomly. First of all, the set of demand points are generated according to a uniform distribution as proposed by Codina, E. (2012). We determine demand values as the Euclidean distance between the origin destination points.

Then, demand realizations at each node $i \in \mathcal{N}$, $d \in \mathcal{D}$ are generated to see the effects of the CVaR value of the expectation model optimum solution with the CVaR value of our model. 5 different significant demand distributions are considered such as [50, 150], [50, 200], [100, 150], [100, 200], and [150, 200] with uniformly distributed lower and upper values respectively. The demand should be between the given intervals.

Different than demand, time realizations are generated with the following (5.1),

$$t^s = \bar{t} + unif[0, 2.5]\bar{t} \tag{5.1}$$

5 different meaningful time distributions are related with the following values, [0, 0.5], [0, 1], [0, 1.5], [0, 2], and [0, 2.5] with uniformly distributed lower and upper values respectively. For [0, 2.5] time generation, according to (5.1), the time can be increased up to %250 of the original time value that indicates the "2.5" value in the uniform distribution, or the original time value will not be changed which is the "0" value in the uniform distribution.

The maximum number of vehicles that can be used in the transit network, M is selected from the uniform distribution with the interval [0, 8]. As "0" represents minimum number of vehicles, and "8" represents maximum number of vehicles in the network. After a variety of trials on the model, the most significant solutions for the fleet size are determined such as, 2, 2.5, 3, 3.5, and 4.

Scenario probabilities $p_s, s \in \mathcal{S}$, are set to be equal. In this study, 10 scenarios, 30 scenarios, 50 scenarios, and 100 scenarios are considered as scenario based model. To compare all scenarios under same conditions we use these 4 different scenarios. With these different scenarios one can see explicitly the differences on overall time on each stop which includes waiting time, and in-vehicle time on each arc which is related with the demand of each node.

In the presented study, we also focus on the α value which is associated with the CVaR realization formula. After several tests on the model, 3 different most effective confidence level values are noted to compare the scenarios. We use 0.7, 0.8, and 0.9 as confidence level values to show the effects for each scenario.

5.2 Computational effectiveness

We provide results to demonstrate the computational effectiveness of the proposed CVaR model.

To see the significance results of the solution the fleet size that is M value is focused separately for each demand and time values. For instance, for the fleet size $M = 2$, all significant demand distributions, and all meaningful time distributions which are selected form several trials are considered together. The model compares the CVaR value of the expectation model optimum solution with the CVaR value of our model. That is why one can see clearly that the results of our model gives better results in CVaR under higher uncertainty such as in bigger fleet size, and under higher uniform distribution on demand such as [150, 200] with lower and upper values respectively, and under higher time such as [0, 2.5] with lower and upper values respectively.

To see the significant results of the solution, the fleet size M is considered for each demand and time values. For instance, for the fleet size $M = 2$, all significant demand distributions, and all critical time distributions which are selected form several trials, are considered together. The model compares the CVaR value of the expectation model optimum solution with the CVaR value of our model. That is why one can see clearly that the results of our model gives better results in CVaR under higher uncertainty such as in bigger fleet size, and under higher uniform distribution on demand such as [150, 200] with lower and upper values respectively, and under higher time such as [1.0, 2.0] with lower and upper values respectively.

The demand matrix is given for all scenarios from node 1 to node 3. That means the demand can be transferred in two ways. One of them is, demand can be transferred from 1 directly to destination node 3 by using the arc (1,3), and the other transfer is from node 1 through node 2, then to node 3 by using the arcs (1,2) and (2,3) respectively. The time for arc (1,3) is 0.15 hours, for arc (1,2) is 0.30 hours, and for arc (2,3) is 0.15 hours. For each scenario, the time of the arcs are constant.

Table 5.1: Time on each arc for 10 scenarios related with the toy problem (in hours).

		For scenario 10									
		1	2	3	4	5	6	7	8	9	10
Arc	(1,3)	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
	(1,2)	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
	(2,3)	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15

Table 5.2: Changing time after applying uniform distribution (in hours).

		For scenario 10									
		1	2	3	4	5	6	7	8	9	10
Arc	(1,3)	0.362	0.150	0.241	0.386	0.383	0.150	0.150	0.352	0.329	0.150
	(1,2)	0.300	0.300	0.300	0.300	0.300	0.80	0.300	0.300	0.300	0.300
	(2,3)	0.150	0.447	0.150	0.150	0.150	0.150	0.280	0.150	0.150	0.250

Table 5. 2 represents the generated time values. The above table defines the uniformly distribution time with lower level "0" where the time value will not be changed, and with upper level "1.5" where the time can be increased up to % 150 of the original time value.

CVaR value of the expectation model optimum solution with the CVaR value of our model is compared. $fcvarexp$ denotes the value of the expectation model that is solved with CVaR objective function. $fexpexp$ denotes the value of the expectation model that is solved with the objective function of the expectation model. Then, we define the CVaR value of the expectation model optimum solution as follows:

$$\frac{fcvarexp - fexpexp}{fexpexp} \tag{5.2}$$

$fexpcvar$ denotes the optimum solution of our model that is solved with the objective function of the expectation model. $fcvarcvar$ denotes the optimum solution of our model that is solved with CVaR objective function. Then, we define the CVaR value of our model optimum solution as follows:

$$\frac{fexpcvar - fcvarcvar}{fexpcvar} \tag{5.3}$$

The first part of Table 5.3 shows the optimum solutions that are obtained from expectation model, that is shown as EXP in the Table 5.3. The second part of Table 5.3 shows the optimum solutions that are obtained from our model, that is shown as CVaR in the Table 5.3.

The performance of EXP and CVaR models are considered in the Table 5.3 when $M = 2$ on determined demand and time values which are uniformly distributed between the values given in the table. The generation demand and time values are mentioned in Section 5.1.

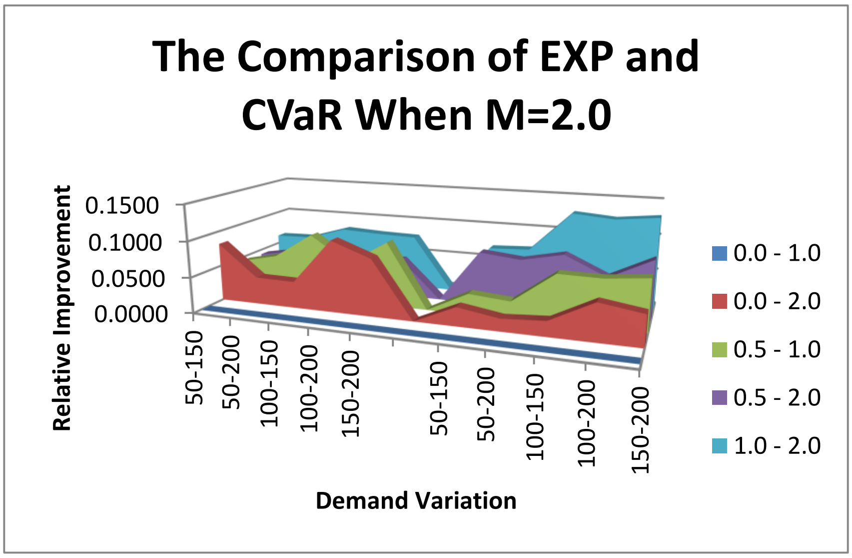

Figure 5.1: The Comparison of EXP and CVaR When M=2.0

As shown in Figure 5.1 demonstrate the demand and time variation on the EXP and CVaR models when the fleet size is 2. The right side area shows the relative improvement of CVaR model and left side area shows the relative improvement of EXP model. For higher time and demand variation CVaR model gives higher improvements.

Table 5.3: The performance of EXP and CVaR models when M = 2 for defined demand and time values

		EXP M = 2.0					**CVaR M = 2.0**				
		demand					**demand**				
		50-150	**50-200**	**100-150**	**100-200**	**150-200**	**50-150**	**50-200**	**100-150**	**100-200**	**150-200**
time	**0.0 - 1.0**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0 - 2.0	0.0821	0.0382	0.0374	0.1004	0.0800	0.0263	0.0174	0.0211	0.0514	0.0428
	0.5 - 1.0	0.0449	0.0572	0.0925	0.0569	0.0907	0.0290	0.0246	0.0678	0.0670	0.0729
	0.5 - 2.0	0.0430	0.0526	0.0678	0.0483	0.0538	0.0745	0.0687	0.0809	0.0548	0.0839
	1.0 - 2.0	0.0622	0.0635	0.0804	0.0775	0.0762	0.0688	0.0715	0.1277	0.1230	0.1293

As one can see that for the same given data when fleet size M is equal to 2, the expected cost value has highest value on uniformly distributed demand between 100 and 200, and also with the uniformly distributed time between the value 0.0 and 1.0. As the 0.0 on time mention no changes of the original time value, and 1.0 mention % 100 additional increases is available on the original time value. On the other hand, when M is equal to 2, the CVaR value has highest value on uniformly distributed demand between 100 and 150, and also with the uniformly distributed time between the value 1.0 and 2.0. That means CVaR value is more sensible while the overall time is increasing %200 percent of the original time value. As our objective function is minimizing overall time that includes both waiting time of the vehicle and in-vehicle travel time, objective function of CVaR gives better results for the network in Figure 5.1.

Moreover, we focus on different number of vehicle size as we mentioned in Section 5.1, such as *M = 2.5, M = 3, M = 3.5, M = 4.* To remember one more time, the vehicle size is determined after various trials on our model. For each vehicle size, the details on selected demand and time values with results are given.

In table 5.4, we can see the relative improvements for each fleet size and for each time interval with a constant demand variation. This demand variation value is selected after several trials and most effective demand is selected to see the effects on fleet size and time variation at the same time.

Figure 5.2 shows the travel time variation and fleet size variation of the model. If the travel time variation is too high or low, or if the fleet is too small or large, the improvement decreases. For example in Table 5.4, with maximum fleet size 3, maximum improvement occurs in time interval [0.5, 1.0] with %14. However, with the same fleet size the improvement can be seen as %5.44 between minimum time interval [0.0, 1.0] and maximum time interval [1.0, 2.0]. On the other hand, when we consider minimum fleet size 2, maximum improvement occurs with minimum time interval [0.0, 1.0] as %9.36, and minimum improvement occurs with time interval [0.5, 2.0] as

%1.15. Higher improvement occurs with %14.00 when the fleet size $M = 3$ and time variation is [0.5, 1.0] for demand between 100 and 150.

Table 5.4: Time variation vs. fleet size variation with a constant demand interval for CVaR models

	demand 100 - 150				
	0.0 - 1.0	**0.0 - 2.0**	**0.5 - 1.0**	**0.5 - 2.0**	**1.0 - 2.0**
M = 2	9.36%	3.51%	3.59%	1.15%	2.04%
M = 2.5	11.77%	5.57%	12.60%	5.22%	6.74%
M = 3	9.65%	5.73%	14.00%	7.31%	5.44%
M = 3.5	8.19%	5.56%	12.32%	6.62%	7.83%
M = 4	6.90%	5.27%	10.90%	5.47%	7.68%

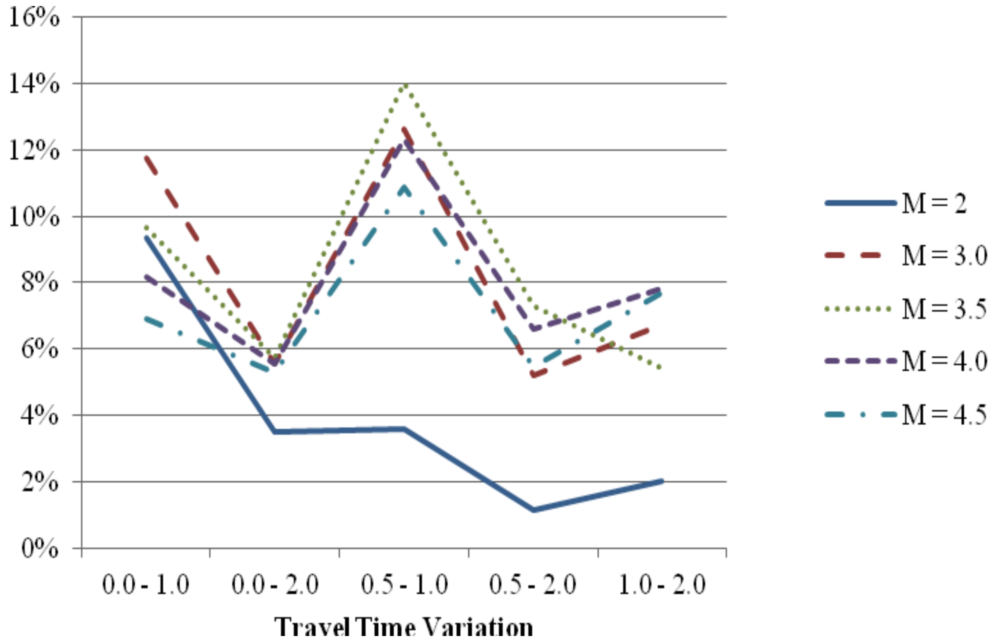

Figure 5.2: The effects of travel time variation and fleet size variation

In table 5.5, we can see the relative improvements for each fleet size and for each demand variation with a constant time interval. This time interval is selected after several trials and most effective time interval is selected to see the effects on fleet size and demand variation simultaneously.

For example in Table 5.5, maximum improvement occurs when fleet size *M* is 2.5 with demand variation [150, 200] for time interval [0.0, 1.5] with 13.89%. But, when we consider minimum fleet size *M* as 2, the improvement is around 3.5% that is too small for all demand variations. Also, when we consider maximum fleet size 4, maximum improvement can be seen around 10.5%. Thus, the higher improvement occurs around 14.00% for all demand variations.

Table 5.5: Demand variation vs. fleet size variation with a constant time interval for CVaR models

	time 0.0-1.5				
	50-150	**50-200**	**100-150**	**100-200**	**150-200**
M = 2	3.62%	3.59%	3.41%	3.52%	3.40%
M = 2.5	10.09%	12.60%	12.37%	12.59%	13.89%
M = 3	12.57%	14.00%	13.42%	13.48%	13.48%
M = 3.5	10.85%	12.32%	11.67%	11.73%	11.77%
M = 4	9.38%	10.90%	10.18%	10.25%	10.32%

The demand variation and fleet size variation of the model is shown in Figure 5.3. As the vehicle capacity is not considered, demand variation does not significantly affect improvement. If the fleet is too small or too large, the improvement decreases.

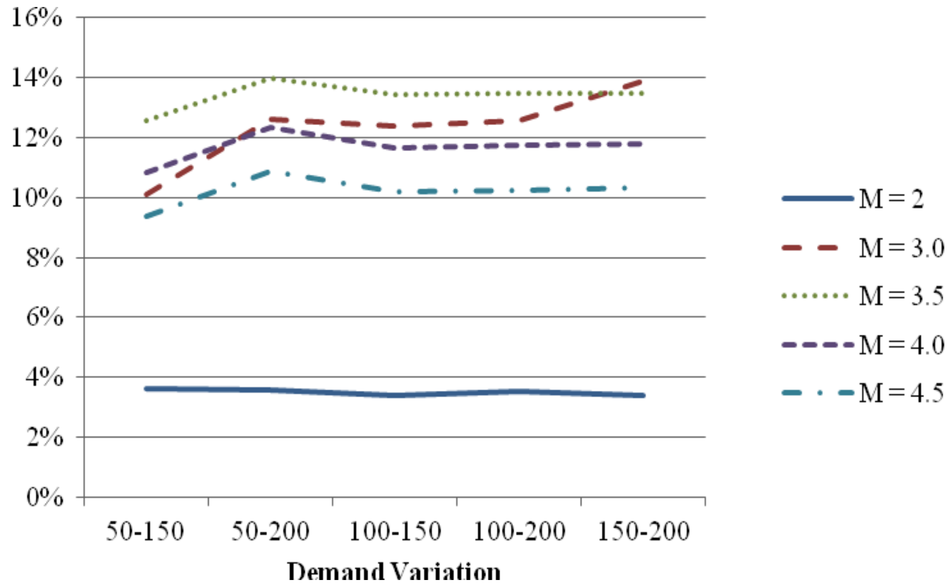

Figure 5.3: The effects of demand variation and fleet size variation

In Table 5.6, we consider fleet size as 3.5 to see the demand variation and time variation effects at the same time, according to the improvement results of the Table 5.4 and Table 5.5. When the fleet size is 3.5, maximum improvement occurs with 14.00% in time interval [0.5, 1.0] and demand variation [50, 200]. Also, high improvements occur for all demand variation around 13.50% in time interval [0.5, 1.0]. Also, minimum improvements occur around 6% when time variation is between [0.0, 2.0] for all demand variations.

Travel time variation and demand variation of the model is presented in Figure 5.4. If the travel time variation is too high or low, the improvement decreases. The relative improvements of Table 5.6 can be seen explicitly in Figure 5.4. Maximum improvement is around 14.00% when time interval is between [0.5, 1.0], and minimum improvement is around 6% when time interval is between [0.0, 2.0]. When we don't change the time as in equation (5.1) $t^s = \bar{t} + unif(0,3)\bar{t}$, the improvement is not effected too much. Thus, it is clear to see minimum improvement around 6% when time interval is between [0.0, 2.0].

Table 5.6: Demand variation vs. time variation with a constant fleet size for CVaR models

		M=3.5				
		demand				
		50-150	**50-200**	**100-150**	**100-200**	**150-200**
time	**0.0 - 1.0**	9.69%	9.65%	8.87%	10.18%	9.37%
	0.0 - 2.0	7.46%	5.73%	4.17%	5.92%	5.86%
	0.5 - 1.0	12.57%	14.00%	13.42%	13.48%	13.48%
	0.5 - 2.0	5.79%	7.31%	7.02%	6.85%	6.98%
	1.0 - 2.0	7.08%	5.44%	6.87%	6.91%	8.03%

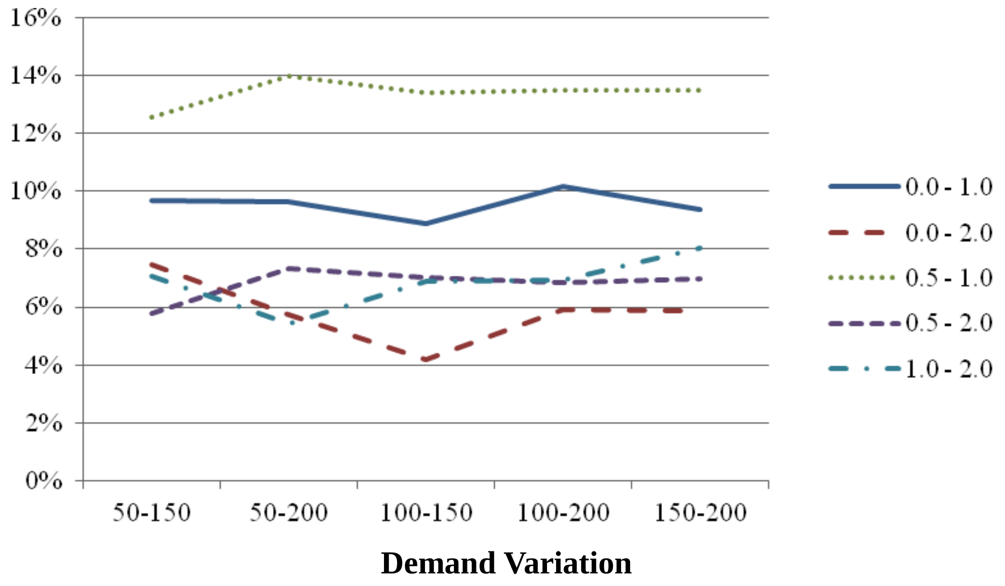

Figure 5.4: The effects of demand variation and travel time variation

Table 5.7 shows the scenario variations for α values are 0.7, 0.8, and 0.9. For each scenario numbers, when α value is increased the relative improvement of the model is increased. Also, for each α value, when scenario numbers are increased the relative improvement of the model is also increased. Maximum improvement occurs with 25% when α value is 0.9 and when scenario number is 100. Minimum improvement occurs around 12% for all scenarios with 0.7 α value.

Table 5.7: Scenario variation vs. alpha variation for CVaR models

		CVaR			
		scenario			
		10	**30**	**50**	**100**
alpha	**0.7**	0.0915	0.0654	0.1527	0.1811
	0.8	0.1209	0.0901	0.1893	0.2159
	0.9	0.1466	0.1096	0.2206	0.2499

Figure 5.5 presents the scenario variation and α value variation of the model. In general, the improvement increases when the number of scenarios and/or α value also increase.

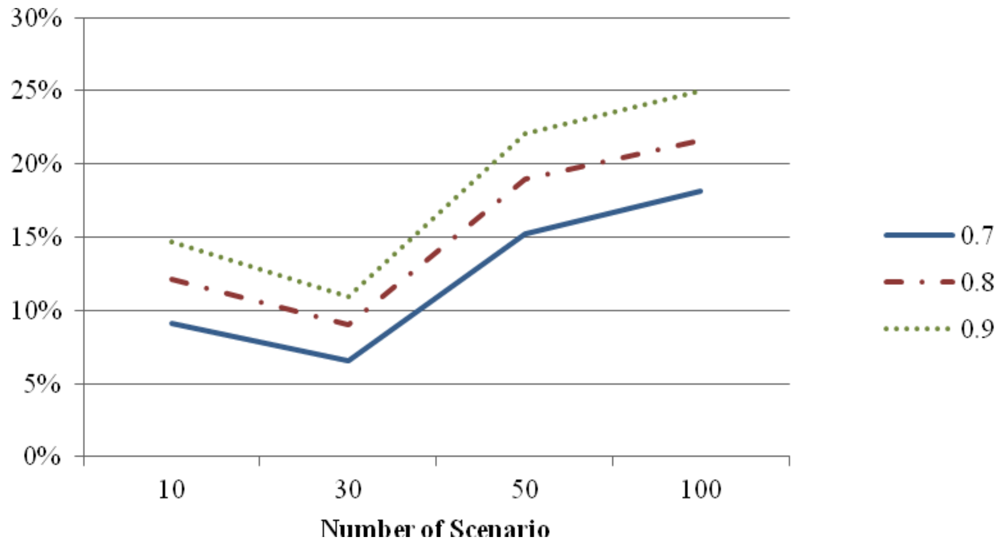

Figure 5.5: The effects of number of scenarios and α value

In addition, as it has been noted before we use different α levels such as 0.7, 0.8, and 0.9, and different number of scenarios such as 10, 30, 50, and 100 to check the reliability of our model, simultaneously. The CVaR value of our model gives better values than CVaR value of the expectation model optimum solution. Therefore, we recommend using CVaR to determine transit frequencies that achieve more accurate results.

Here, the comparison between expectation model and CVaR model is given in Table 5.8. In Figure 5.6, the area on the right side is the relative improvement of CVaR model which is higher than the expectation model.

Table 5.8: The effect of Confidence Level values

		EXP				**CVaR**			
		scenario				**scenario**			
		10	**30**	**50**	**100**	**10**	**30**	**50**	**100**
α	**0.7**	0.0255	0.0404	0.0460	0.0510	0.0206	0.0423	0.0316	0.0297
	0.8	0.0418	0.0670	0.0815	0.0793	0.0380	0.0750	0.0727	0.0717
	0.9	0.0852	0.0878	0.0979	0.0889	0.1016	0.1225	0.1260	0.1339

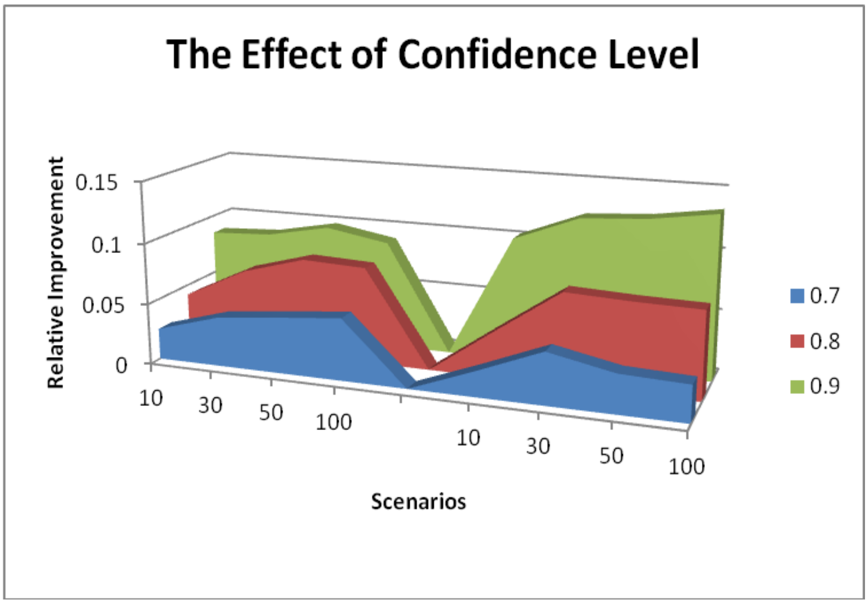

Figure 5.6: The effects of Confidence Level values

Table 5.9: The performance of EXP and CVaR models when M = 2.5 for defined demand and time values

		EXP M = 2.5					CVAR M = 2.5				
		demand					**demand**				
		50-150	**50-200**	**100-150**	**100-200**	**150-200**	**50-150**	**50-200**	**100-150**	**100-200**	**150-200**
time	**0.0 - 1.0**	0.0098	0.0081	0.0162	0.0000	0.0098	0.0018	0.0003	0.0004	0.0000	0.0018
	0.0 - 2.0	0.0697	0.0403	0.0515	0.1074	0.1070	0.0444	0.0257	0.0360	0.0714	0.0652
	0.5 - 1.0	0.0537	0.0574	0.0794	0.0577	0.0820	0.0441	0.0425	0.0924	0.0686	0.0955
	0.5 - 2.0	0.0481	0.0573	0.0726	0.0522	0.0614	0.0824	0.0738	0.1077	0.0757	0.1059
	1.0 - 2.0	0.0676	0.0611	0.0747	0.0717	0.0715	0.0769	0.0879	0.1297	0.1230	0.1191

The first part of Table 5.9 shows the optimum solutions that are obtained from expectation model, that is shown as EXP in the Table 5.3. The second part of Table 5.9 shows the optimum solutions that are obtained from our model, that is shown as CVaR in the Table 5.9.

In Figure 5.7 the comparison between expectation model and CVaR model can be seen better when the fleet size is increased to 2.5 and as the demand variation is increased CVaR has better results.

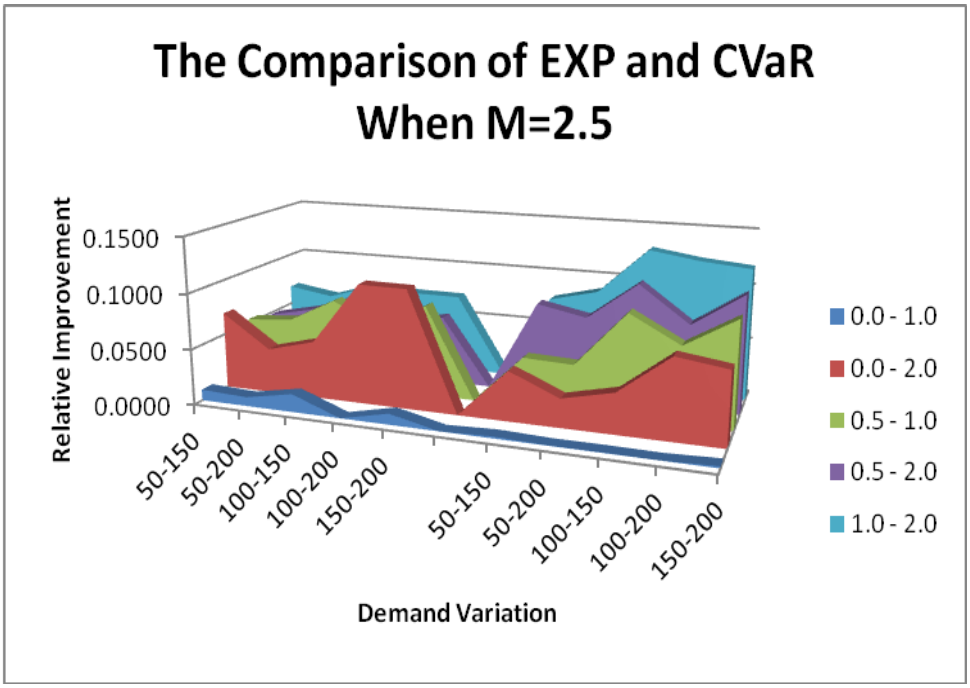

Figure 5.7: The Comparison of EXP and CVaR When M=2.5

Table 5.10: The performance of EXP and CVaR models when M = 3 for defined demand and time values

		EXP M = 3.0					CVAR M = 3.0				
		demand					demand				
		50-150	**50-200**	**100-150**	**100-200**	**150-200**	**50-150**	**50-200**	**100-150**	**100-200**	**150-200**
time	**0.0 - 1.0**	0.0368	0.0066	0.0234	0.0095	0.0239	0.0037	0.0019	0.0045	0.0010	0.0041
	0.0 - 2.0	0.0654	0.0487	0.0803	0.0933	0.1014	0.0594	0.0353	0.0411	0.0911	0.0808
	0.5 - 1.0	0.0515	0.0743	0.0692	0.0579	0.0725	0.0570	0.0582	0.1115	0.0801	0.1127
	0.5 - 2.0	0.0468	0.0520	0.0630	0.0415	0.0502	0.0916	0.0940	0.0959	0.0827	0.1131
	1.0 - 2.0	0.0584	0.0552	0.0726	0.0599	0.0616	0.0771	0.0861	0.1120	0.1187	0.1191

In Table 5.10, the fleet size is increased to 3. Although expectation model gives sufficient results when time variation is between [0.0-2.0], CVaR is more reliable to see the effects when demand variation, time variation and fleet size variations are increased.

In Figure 5.8 the comparison between expectation model and CVaR model can be seen better when the fleet size is increased. Also, it can be seen that as the demand variation is increased CVaR gives better results than expectation model.

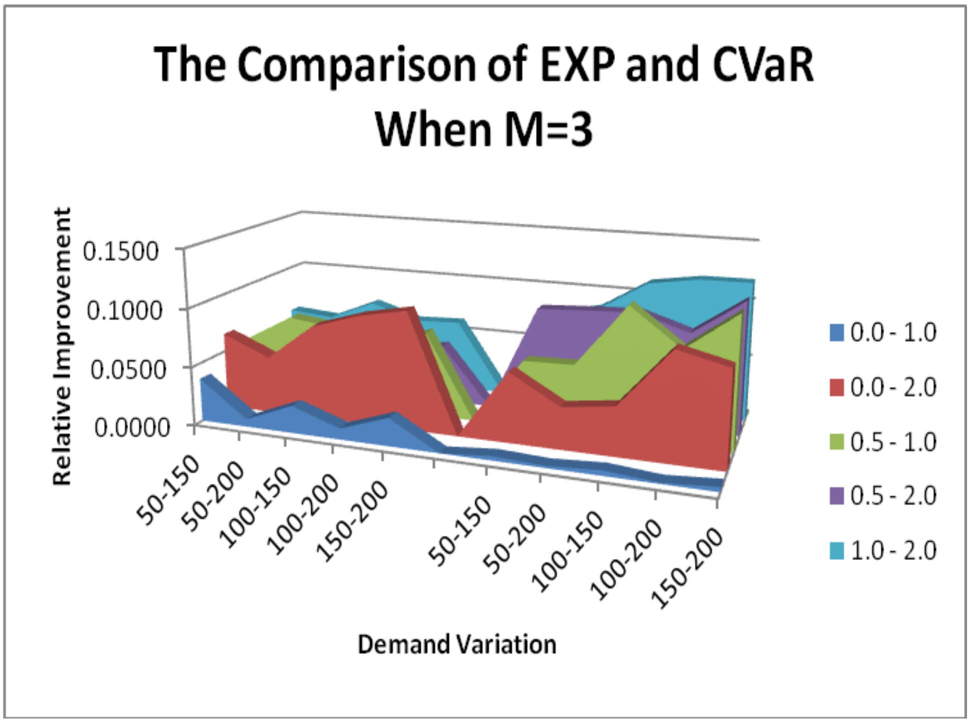

Figure 5.8: The Comparison of EXP and CVaR When M=3

Table 5.11: The performance of EXP and CVaR models when M = 3.5 for defined demand and time values

		EXP M = 3.5					CVAR M = 3.5				
		demand					demand				
		50-150	50-200	100-150	100-200	150-200	50-150	50-200	100-150	100-200	150-200
time	0.0 - 1.0	0.0367	0.0211	0.0203	0.0179	0.0391	0.0076	0.0033	0.0080	0.0036	0.0076
	0.0 - 2.0	0.0655	0.0477	0.0794	0.1009	0.1067	0.0721	0.0456	0.0565	0.0984	0.0963
	0.5 - 1.0	0.0546	0.0654	0.0599	0.0620	0.0719	0.0679	0.0715	0.1128	0.0933	0.1273
	0.5 - 2.0	0.0388	0.0418	0.0572	0.0338	0.0431	0.0924	0.0952	0.0960	0.0880	0.0994
	1.0 - 2.0	0.0518	0.0559	0.0620	0.0522	0.0531	0.0802	0.0939	0.1287	0.1103	0.1333

In Table 5.11, the fleet size is increased to 3.5. Here, it is obviously shown that the CVaR model has better results than expectation model. Almost %13 improvement is obtained by using CVaR model. This is obtained when the time variation is between [1.0-2.0] and the demand variation is between [100-150] with the fleet size 3.5.

In Figure 5.9, it is explicitly shown that the CVaR model has better results when the worst cases are occurred. Thus, the CVaR model gives better results than expectation model as the demand variation is increased.

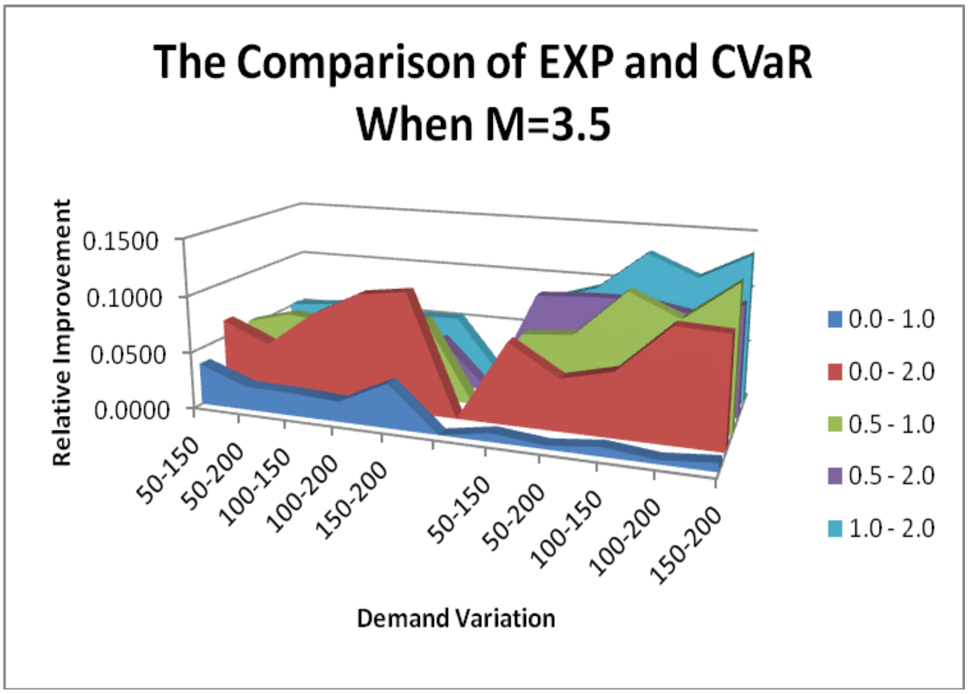

Figure 5.9: The Comparison of EXP and CVaR When M=3.5

Table 5.12: The performance of EXP and CVaR models when M = 4 for defined demand and time values

		EXP M = 4.0					CVAR M = 4.0				
		demand					demand				
		50-150	**50-200**	**100-150**	**100-200**	**150-200**	**50-150**	**50-200**	**100-150**	**100-200**	**150-200**
time	**0.0 - 1.0**	0.0362	0.0179	0.0339	0.0394	0.0560	0.0185	0.0086	0.0186	0.0133	0.0225
	0.0 - 2.0	0.0582	0.0614	0.0725	0.0886	0.0936	0.0704	0.0544	0.0639	0.1040	0.0987
	0.5 - 1.0	0.0579	0.0693	0.0679	0.0677	0.0624	0.0729	0.0832	0.1289	0.0934	0.1417
	0.5 - 2.0	0.0381	0.0351	0.0559	0.0284	0.0371	0.0786	0.0938	0.0957	0.0568	0.0844
	1.0 - 2.0	0.0457	0.0507	0.0523	0.0503	0.0473	0.0644	0.0809	0.1119	0.1218	0.1169

In Table 5.12, the fleet size is increased to 4. Here, it is obviously shown that the CVaR model has better results than expectation model. Almost %15 improvement is obtained by using CVaR model. This is obtained when the time variation is between [0.5-1.0] and the demand variation is between [100-150] with the fleet size 4.

In Figure 5.10, it is explicitly shown that the CVaR model has better results when the worst cases are occurred. Thus, the CVaR model gives better results than expectation model as the demand variation and time variation is increased.

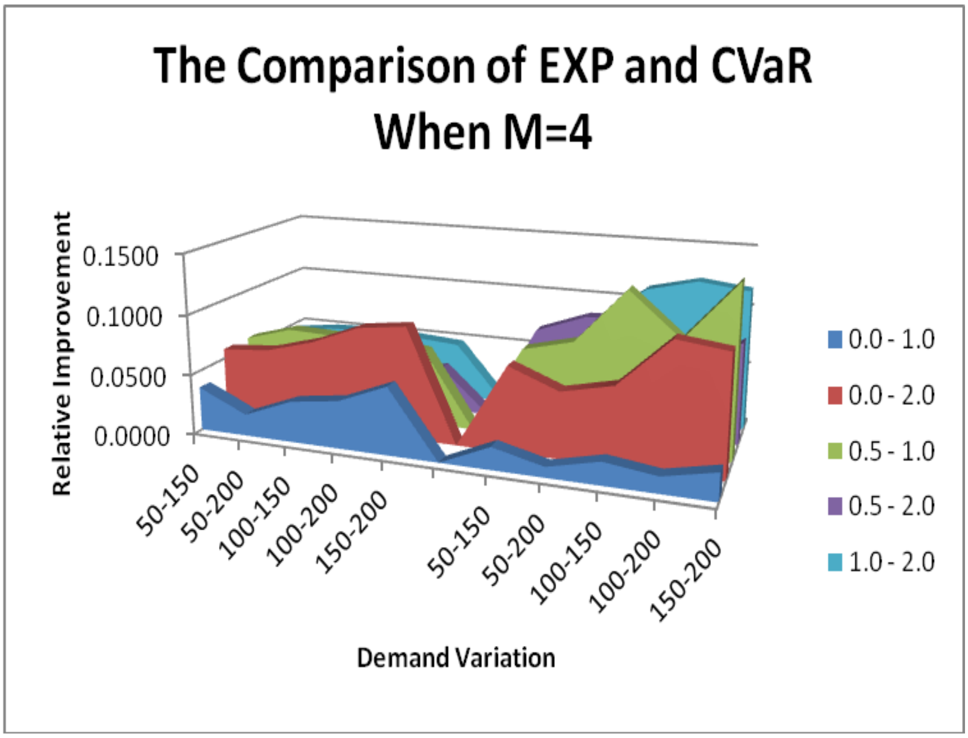

Figure 5.10: The Comparison of EXP and CVaR When M=4

6 CONCLUSION

With the increasing number of registered vehicles in traffic, people are spending too much time in traffic to go from their home to their jobs, schools, etc. Due to the traffic congestion, passengers total travel time is increased. By decreasing private car ownership, people need to be redirected to the public transportation system.

Transit network frequency setting is considered as the minimization of total in-vehicle and at stop waiting times. While frequencies of transit system lines are planned, minimization of the total travel time spent by the passengers is the most preferred objective. Generally travel time, waiting time, etc.. is determined by using the mean or expected value. However, the determined line frequency based on these data cause passengers to travel in the excessing durations, especially with the variability of traffic congestion.

As the travelling conditions may change unfavorably and the existing transit systems are planned with traditional expected time approaches, passengers occasionally take longer times to travel. Unlike from other studies in the literature, a risk averse approach is adopted in this study to find line frequencies. Optimum line frequencies are identified by considering stochastic travel times and demands which are due to uncertain network conditions. Moreover, random parameters are represented with finite number of scenarios. Conditional value-at-risk (CVaR) measure is used as a risk modeling approach. We obtained important improvements on the total traveling times of passengers with our approach compared to the expected value based approach under uncertain traveling conditions.

This study has the potential of being a starting point for many future researches. We can only conceive of apparent ones. As for example, the capacity of the vehicles and the comfort of passenegers (utility) during their trip may be included to the model.

Besides, network design and/or timetabling stages may be integrated to the frequency setting stage. Risk measure other than CVaR can be used. Also, different scenario generation methods can be developed. Instead of single node, multiple node can be deteriorated by changing demand and travel time. Consecutive links can be used to create a different type of scenario. Further, the validity of the approach can be tested with real transit network data.

References

Aggarwal, J., Mathew, T.V. (2004). Transit Route Network Design Using Genetic Algorithms. Journal of Computing in Civil Engineering, 18, p.248-256.

Anandalingam, G., Friesz, T. (1992). Hierarchical optimization: an introduction. Annals of Operations Research, 34, p.1–11.

Artzner, P., Delbaen, F., Eber, J., Heath, D. (1999). Coherent measures of risk. Mathematical Finance, 9(3), p.203-228.

Automotive Industry Association. Monthly Report. (2013). URL: http://www.osd.org.tr/2013rapor.pdf [accessed January 02, 2014]

Azad, N. & Davoudpour, H. (2013). Designing a stochastic distribution network model under risk. The International Journal of Advanced Manufacturing Technology, 64, p.23–40.

Bell, M., G., H., Cassir, C. (2002). Risk-averse user equilibrium traffic assignment: an application of game theory. Transportation Research Part B, 36, p.671–681.

Ben-Ayed, O. (1993). Bilevel linear programming. Computers and Operations Research, 20, p.485–501.

Ben-Ayed, O., & Blair, C. (1990). Computational difficulties of bilevel linear programming. Operations Research, 38, p.556–560.

Calvete, H.I., Galé, C. & Mateo, P.M. (2008). A New Approach for Solving Linear Bilevel Problems Using Genetic Algorithms. European Journal of Operational Research, 188(1), p.14-28.

Ceder, A., Chowdhury, S., Taghipouran, N., Olsen, J. (2013). Modelling public-transport users' behaviour at connection point. Transport Policy, 27, p.112–122.

Cepeda, M., Cominetti, R. & Florian, M. (2006). A Frequency-Based Assignment Model For Congested Transit Networks With Strict Capacity Constraints: Characterization And Computation Of Equilibria. Transportation Research Part B: Methodological, 40(6), p.437–459.

Ceylan, H., Bell, M.G.H. (2004). Traffic Signal Timing Optimization Based on Genetic Algorithm Approach Including Drivers' Routing. Transportation Research Part B: Methodological, 38, p.329-342.

Chen, A., Zhou, Z. (2010). The a-reliable mean-excess traffic equilibrium model with stochastic travel times. Transportation Research Part B, 44, p.493–513.

Chen, A., Zhou, Z., Chootinan, P., Ryu, S., Yang, C., Wong, S., C. (2011). Transport Network Design Problem under Uncertainty: A Review and New Developments. Transport Reviews: A Transnational Transdisciplinary Journal, 31(6), p.743-768.

Choi, P. & Min, I. (2011). A Comparison of Conditional and Unconditional Approaches in Value-At-Risk Estimation. Japanese Economic Review, 62(1), p.99–115.

Codina, E. (2012). A Variational Inequality Reformulation of a Congested Transit Assignment Model by Cominetti, Correa, Cepeda, and Florian.Transportation Science, Articles in Advance, p.1–16.

Colson, B., Marcotte, P., Savard, G. (2007). An overview of bilevel optimization. Annals of Operation Research, 15(1), p.235-256.

Cominetti, R., Correa, J. (2001). Common-Lines and Passenger Assignment in Congested Transit Networks. Transportation Science, 35(3), p.250–267.

Constantin, I., Florian, M. (1995). Optimizing Frequencies In A Transit Network: A Nonlinear Bi-Level Programming Approach. International Transactions in Operational Research, 2(2), p.149–164.

Côté, J.P., Marcotte, P., Savard, G. (2003). A bilevel modeling approach to pricing and fare optimization in the airline industry. Journal of Revenue and Pricing Management, 2, p.23–36.

Deb, K. & Sinha, S. (2010). An Efficient and Accurate Solution Methodology for Bi-Level Multi-Objective Programming Problems Using a Hybrit Evalutionary-Local-Search Algorithm. Journal Evolutionary Computation, 18(3), p.403-449.

Dempe, S. (2003). Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization, 52, p.333–359.

Desaulniers, G., Hickman, M. (2007). Public transit. Handbooks in Operation Research and Management Science, p.69–120.

Dimitriou, L., Stathopoulos, A. (2008). Reliable Stochastic Design of Road Network Systems. International Journal of Industrial and Systems Engineering, 3, p.549-574.

Dimitriou, L., Tsekeris, T., Stathopoulos, A. (2008). Applications of Evalutionary Computing, Chapter Genetic Computation of Road Network Design and Pricing Stackelberg Games with Multi-Class Users, p.669-678. Springer, Berlin/Heidelberg.

Drezner, Z., Salhi, S. (2002). Using Hybrid Metaheuristics for the One-Way and Two-Way Network Design Problem. Naval Research Logistics, 49, p.449-463.

Drezner, Z., Wesolowsky, G.O. (2003). Network Design: Selection and Design of Links and Facility Location. Transportation Research Part A: Policy and Practice, 37, p.241-256.

Embrechts, P., Kuuppelberg, C., Mikosch, T. (1997). Extremal Events in Finance and Insurance. Springer, Berlin.

Florian, M. (ed) (2003). Frequency Based Transit Route Choice Models. Advanced Modeling for Transit Operations and Service Planning, Pergamon, p.165-180.

Florian, M., Chen, Y. (1995). A coordinate descent method for the bi-level o-d matrix adjustment problem. International Transactions in Operational Research, 2, p.165–179.

Gallo, M., Montella, B., D'Acierno, L. (2011). The Transit Network Design Problem with Elastic Demand and Internalisation of External Costs: An Application to Rail Frequency Optimisation. Transportation Research Part C: Emerging Technologies, 19(6), p.1276–1305.

Gardner, L.M., Unnikrishnan, A., Waller, S.T. (2010). Solution Methods for Robust Pricing of Transportation Networks Under Uncertain Demand. Transportation Research Part C: Emerging Technologies, 18(5), p.656-667.

Guihaire, V. & Hao, J.-K. (2008). Transit Network Design And Scheduling: A Global Review. Transportation Research Part A: Policy and Practice, 42(10), p.1251–1273.

Hadas, Y., Shnaiderman, M. (2012). Public-Transit Frequency Setting Using Minimum-Cost Approach with Stochastic Demand and Travel Time. Transportation Research Part B: Methodological, 46(8), p.1068–1084.

Hamdouch, Y. Et Al. (2011). Schedule-Based Transit Assignment Model with Vehicle Capacity and Seat Availability. Transportation Research Part B: Methodological, 45(10), p.1805–1830.

Hearn, D.W., Romana, M.V. (1998). Solving Congestion Toll Pricing Models. In Marcotte, P., Nguyen, S. Editors, Equilibrium and Advanced Transportation Modeling, p.109-124. Kluwer Academic Publishers, Dordrecht.

Hobbs, B. F., Nelson, S. K. (1992). A nonlinear bilevel model for analysis of electric utility demand-side planning issues. Annals of Operations Research, 34, p.255–274.

Hsu, S., Wen, U. (1989). A review of linear bilevel programming problems. In Proceedings of the National Science Council, Republic of China, Part A: Physical Science and Engineering, 13, p.53–61.

Kara, B. Y., Verter, V. (2004). Designing a road network for hazardous materials transportation. Transportation Science, 38, p.188–196.

Karoonsoontawong, A., Waller, S.T. (2006). Dynamic Continuous Network Design Problem: Linear Bilevel Programming and Metaheuristic Approaches. Journal of Transportation Research Board, 1964, p.104-117.

Know, C. (2011). Conditional Value-At-Risk Model for Hazardous Materials Transportation. Proceedings of the 2011 Winter Simulation Conference, p.1708–1714.

Kuo, R.J, Huang, C.C. (2009). Application of Particle Swarm Optimization Algorithm for Solving Bi-Level Linear Programming Problem. Computers & Mathematichs with Applications, 58(4), p.678-685.

Lam, W.H.K., Zhou, J., Sheng, Z. (2002). A Capacity Restraint Transit Assignment with Elastic Line Frequency. Transportation Research Part B: Methodological, 36(10), p.919–938.

Leiva, C. Et Al. (2010). Design Of Limited-Stop Services for an Urban Bus Corridor with Capacity Constraints. Transportation Research Part B: Methodological, 44(10), p.1186–1201.

Lim, A.E.B., Shanthikumar, J.G., Vahn, G.-Y. (2011). Conditional Value-At-Risk in Portfolio Optimization: Coherent But Fragile. Operations Research Letters, 39(3), p.163–171.

Lo, H.K., Luo, X.W., Siu, B.W.Y. (2006). Degradable transport network: Travel time budget of travelers with heterogeneous risk aversion. Transportation Research Part B, 40, p.792–806.

Mathew, T.V., Shrama, S. (2009). Capacity Expansion Problem for Large Urban Transportation Networks. Journal of Transportation Engineering, 135, p.406-415.

Monyrath, K., 2006. Frequency Design Of Bus Transit In Mixed Urban Transport Network. , p.1–6.
URL: http://www.enveng.titech.ac.jp/yai/pdf/2006/kov200703.pdf [accessed May17, 2013].

Monyrath, K., Fukuda, D., Yai, T. (2011). Optimizing Frequency of Bus Services in Mixed-Traffic Urban Streets. Asian Transport Studies, 1(4), p.352–367.

Marcotte, P., Savard, G. (2005). Bilevel programming: a combinatorial perspective. In: Avis, D., Hertz, A., Marcotte, O. (Eds.), Graph theory and combinatorial optimization. Boston: Kluwer Academic.

Murray, A.T. (2003). A coverage model for improving public transit system accessibility and expanding access. Annals of Operations Research 123, p.143–156.

Rundmo, T., Nordfjærn, T., Iversen, H., H., Oltedal, S., Jørgensen, S., H. (2011). The role of risk perception and other risk-related judgements in transportation mode use. Safety Science, 49, p.226–235.

Schmöcker, J., D., Bell, M., G., H., Kurauchi, F. (2008). A Quasi-Dynamic Capacity Constrained Frequency-Based Transit Assignment Model. Transportation Research Part B: Methodological, 42(10), p.925–945.

Seyedshohadaie, S., R., Damnjanovic, I., Butenko, S. (2010). Risk-based maintenance and rehabilitation decisions for transportation infrastructure networks. Transportation Research Part A, 44, p.236–248.

Shepherd, S.P., Sumalee, A. (2004). A Genetic Algorithm Based Approach to Optimal Toll Level and Location Problem. Netwroks and Spatial Economics, 4, p.161-179.

Sorokin, A., Boginski, V., Nahapetyan, A., Pardalos, P., M. (2013). Computational risk management techniques for fixed charge network flow problems with uncertain arc failures. Journal of Combinatorial Optimization, 25, p.99–122.

Spasovic, L.N., Boile, M.P., Bladikas, A.K. (1993). A methodological framework for optimizing bus transit service coverage. In: Proceedings of the 73rd Annual Meeting of the Transportation Research Board.

Stoyanov, S. V., Rachev, S.T. & Fabozzi, F.J. (2012). Cvar Sensitivity with Respect to Tail Thickness. Journal Of Banking & Finance.
URL: http://Linkinghub.Elsevier.Com/Retrieve/Pii/S0378426612003573 [accessed May17, 2013].

Sumale, A., May, T., Shepherd, S. (2005). Comparison of Judgemental and Optimal Road Pricing Cordons. Transport Policy, 12, p.384-390.

Sumalee, A., Uchida, K., Lam, W., H., K. (2011). Stochastic multi-modal transport network under demand uncertainties and adverse weather condition. Transportation Research Part C, 19, p.338–350.

Szegö, G. (2002). Measures of Risk. Journal of Banking and Finance, 26(2), p.1253-1272.

Szeto, W.Y., O'Brien, L., O'Mahony, M. (2006). Risk-Averse Traffic Assignment with Elastic Demands: NCP Formulation and Solution Method for Assessing Performance Reliability. Networks and Spatial Economics, 6, p.313–332.

Szeto, W.Y., Wu, Y. (2011). A Simultaneous Bus Route Design And Frequency Setting Problem For Tin Shui Wai, Hong Kong. European Journal Of Operational Research, 209(2), p.141–155.

Toumazis, I., Kwon, C. (2013). Routing hazardous materials on time-dependent networks using conditional value-at-risk, Transportation Research Part C, 37, p.73–92.

Teklu, F. (2007). A Stochastic Process Approach for Frequency-Based Transit Assignment with Strict Capacity Constraints. Networks And Spatial Economics, 8(2-3), p.225–240.

Tom, V.M., Mohan, S. (2003). Transit Route Network Design Using Frequency Coded Genetic Algorithm. Transportation Engineering, 129(2), p.186–195.

Turkey Statistical Institute. Newsletter. (2013). URL: http://www.tuik.gov.tr/PreHaberBultenleri.do?id=13533 [accessed January 02, 2014]

UN-HABITAT. Access to transport for the urban poor in Asia. (2009). URL: http://www.unhabitat.org/downloads/docs/7802_90848_GENUS%20ASIA%20workshop.Final%20report.pdf [accessed January 02, 2014]

Uryasev, S. (2000). Conditional Value at Risk: Optimization Algorithms and Applications. Financial Engineering News, 14, p.1-5.

Vicente, L.N., Calamai, P.H. (1994). Bilevel and multilevel programming: a bibliography review. Journal of Global Optimization, 5, p.291–306.

Wang, Y., Li, H., Dang, C. (2011). A New Evalutionary Algorithm for a Class of Nonlinear Bi-Level Programming Problems and Its Global Convergence. INFORMS Journal on Computing, 23(4), p.618-629.

Wen, U., Yang, Y. (1991). Algorithms for solving the mixed integer two-level linear programming problem. Computers and Operations Research, 17, p.133-142.

Xu, T., Wei, H., Hu, G. (2009). Study on Continuous Netwrok Design Problem Using Simulated Annealing and Genetic Algorithm. Expert Systems with Applications, 36, p.1322-1328.

Yau, S. Et Al. (2011). Financial and Operational Decisions in the Electricity Sector: Contract Portfolio Optimization with the Conditional Value-At-Risk Criterion. International Journal Of Production Economics, 134(1), p.67–77.

Yin, Y. (2000). Genetic Algortihms Based Approach for Bilevel Programming Models. Journal of Transportation Engineering, 126, p.115-120.

Yoo, G.-S., Kim, D.-K. & Chon, K.S. (2010). Frequency Design in Urban Transit Networks with Variable Demand: Model and Algorithm. KSCE Journal Of Civil Engineering, 14(3), p.403–411.

Yu, B., Yanga, Z., Suna, X., Yaob, B., Zenga, Q., Jeppesenc, E. (2011). Parallel Genetic Algorithm In Bus Route Headway Optimization. Applied Soft Computing, 11(8), p.5081–5091.

Yu, B., Yang, Z., and Yao, J. (2010). Genetic Algorithm for Bus Frequency Optimization. Journal of Transportation Engineering, 136(6), 576–583.

Zhang, X., Yang, H. (2004). The Optimal Cordon-Based Network Congestion Pricing Problem. Transportation Research Part B: Methodological, 38, p.517-537.

Bibliographical Sketch

İlayda Ülkü was born in September 25, 1987 in İstanbul, Turkey. She finished her high school education in Fatih Girls High School. In 2010, she earned her B.Sc. in Industrial Engineering from İstanbul Kültür University, İstanbul. In 2011, she joined the M.Sc. program in Industrial Engineering of Galatasaray University. She is currently working as a Research Assistant in İstanbul Kültür University since 2010.