

A Demand Driven Airline and Airport Evolution Study

Anand Seshadri

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Civil Engineering

Dr Antonio.A.Trani, Chair

Dr Hojong Baik, Co-Chair

Dr Hanif Sherali

Dr Hesham Rakha

Dr Gerardo Flintsch

May 11th 2009

Blacksburg, VA

Keywords: Schedule Generation, Network Evolution, Passenger Route Choice, Airline Networks, Large-scale system

A Demand Driven Airline and Airport Evolution Study

Anand Seshadri

ABSTRACT

The events of September 11, 2001 followed by the oil price hike and the economic crisis of 2008, have lead to a drop in the demand for air travel. Airlines have attempted to return to profitability by cutting service in certain unattractive routes and airports. Simultaneously, delays and excess demand at a few major hubs have lead to airline introducing service at reliever airports. This dissertation attempts to capture the changes in the airline network by utilizing a supply-demand framework.

ACKNOWLEDGEMENT

I would like to thank my advisor Dr. Antonio Trani, for his valuable guidance and mentorship throughout my stay at Virginia Tech. I would also like to thank Dr. Trani for giving me the chance to work on an exciting and intellectually stimulating research topic.

Next, I would like to thank my committee members, Dr. Hojong Baik, Dr. Hanif Sherali, Dr. Hesham Rakha, and Dr. Gerardo Flintsch for their valuable input and assistance during the course of my research. I would like to thank Mr. Howard Swingle for providing me with research ideas.

Finally, I would like to thank my family for their support and encouragement during the course of my doctorate without which, this dissertation might never have been completed.

Contents

Chapter 1 Introduction.....	1
1.1) Motivation and Purpose	1
1.2) Problem Description.....	3
1.2.1) Air Traffic Network Flow Problem	3
1.2.2) Schedule Generator Problem.....	4
1.2.3) Route Generator Problem.....	4
1.3) Organization of this Document	7
1.4) Contributions of this Study	7
Chapter 2 Literature review	8
2.1) Literature review on the Air Traffic Network Flow Problem.....	8
2.1.1) Entropy Based Formulations	9
2.1.2) Stochastic Choice Based Formulations.....	11
2.1.3) Supply-Demand Based Equilibrium Formulations.....	14
2.1.4) Game-Theoretic Formulations	20
2.1.5) Capacity Constrained Choice Formulations.....	34
2.1.6) Fuzzy Logic Based Formulations	36
2.1.7) Heuristic Based Models	38
2.1.8) Probability Based Formulations	39
2.2) Literature review on the Schedule Generator Problem	42
2.2.1) Mathematical Programming Formulations, Independent of Previous Schedules	42
2.2.2) Mathematical Programming Formulations, Dependent on Previous Schedules	56
2.2.3) Supply-Demand Based Formulations	62
2.2.4) Heuristic Based Formulations	62
2.3) Literature Review on the Route Generator Problem	65
2.3.1) Mathematical Programming Based Formulations	65
2.3.2) Econometric Based Formulations.....	72
2.3.3) Logit Based Formulations	72
2.3.4) Heuristic Based Formulations	73
Chapter 3 Methodology	78
3.1) Introduction.....	78
3.2) The Demand Based-Airline and Airport Evolution	78
3.3) A Feedforward Based Approach to the Demand Based Airline and Airport Evolution.....	79
3.3.1) The Air Traffic (O-T-D) Flow Module.....	82
3.3.2) The Schedule Generator	89
3.3.3) The Fare Prediction Module	91
3.3.4) Network Evolution Submodule	92
Chapter 4 Model Description.....	97
4.1) Introduction.....	97
4.1.1) Model Integration	97
4.2) The Air Traffic Network Flow Problem (O-T-D Model)	99
4.2.1) The Schedule Building Module.....	100
4.2.2) The Network Building Module	103

4.2.3) The Discrete Choice Module.....	104
4.3) Schedule Generator	107
4.3.1) Aircraft Type Classification	108
4.3.2) Stopover Flights.....	108
4.3.3) Model Description	110
4.4) Network Evolver Model	123
4.4.1) Cutting Service on Non-profitable Routes.....	123
4.4.2) Introduction of New Point-to-Point Service	126
4.4.3) Introduction of New Spoke-to-Hub Flights	127
4.4.4) Addition and Removal of Airline Hubs	128
4.4.5) Introduction of Service from New Airports	128
Chapter 5 Model Results and Validation.....	130
5.1) Introduction.....	130
5.2) Sub-Module Results and Validation.....	130
5.2.1) The Air Traffic Network Flow Problem	130
5.2.2) The Schedule Generator.....	141
5.2.3) The Network Evolver.....	158
5.2.4) The Fare Prediction Module.....	161
5.3) Integrated Model Results	165
5.3.1) Mode Choice Trends.....	165
5.3.2) Origin-Transfer-Destination (O-T-D) Trends	165
5.3.3) Network Evolver Trends	166
Chapter 6 Conclusions and Recommendations.....	179
Chapter 7 References.....	184

List of Tables

Table 2.1 Summary of the Methods used for the Air Traffic Network Flow Problem.	41
Table 2.2 Methodologies Used for the Schedule Generator Problem.	64
Table 2.3 Methodologies Used for the Route Generator Problem.	77
Table 3.1 Variation of the ASV Weighting Factor with the Aircraft Mix and Meteorological conditions.	86
Table 3.2 The Demand and Capacity Values for the top 33 airports in the NAS (2006)..	87
Table 3.3 Sample Timetable derived from the Timetable Generator.	91
Table 4.1 Variables Used in the Logit Formulation.	104
Table 4.2 Segmentations Considered in the Logit Formulations.	105
Table 4.3 Hourly Operating Cost and Seats Per Aircraft for Each Aircraft Cluster.	109
Table 4.4 Percentage of Stopover Flights for the Major Airlines in 2004.	110
Table 4.5 Cost per Seat per Hour by Aircraft Type.	118
Table 4.6 Aircraft Operating Costs from the T-100.	125
Table 4.7 Distance Threshold for the Introduction of Point-to-Point Air Service.	127
Table 5.1 Multinomial Logit Results and T-Statistics With the Average Airport Delay for the year 2000.	131
Table 5.2 Multinomial and Nested Logit Results and T-Statistics with the Average Airport Delay for the year 2000.	131
Table 5.3 Multinomial Logit Coefficients for the Year 2006.	132
Table 5.4 Multinomial Logit Model for the Year 2006 with the Fare Variable Dropped.	132
Table 5.5 Multinomial Logit Coefficients for the Direct and Connecting Case With the Travel Time and Number of Legs Variable Dropped.	133
Table 5.6 Coefficients and Hit Rates for the Binary Probit Module to Cut Unprofitable Routes.	159
Table 5.7 Coefficients and Hit Rates for the Binary Probit Module for the Introduction of New Hub-Spoke Connections.	160
Table 5.8 Linear Regression Coefficients for the Fare Trend Model.	164
Table 5.9 Total CA Trips (One-Way).	165
Table 5.10 Total Auto Trips (One-Way).	165
Table 5.11 Total CA Domestic Enplanements (TSAM).	166
Table 5.12 Average Number of Legs Per O-D Pair.	166
Table 5.13 Proportion Of Direct Passengers.	166
Table 5.14 Network Revenue, Costs and Profit.	167
Table 5.15 Number of Segments in the CA Network.	167
Table 5.16 List of Airline Hubs by Year.	168
Table 5.17 List of Airline Hubs by FAA ID.	168
Table 5.18 Partial List of Airports Disconnected from the Network.	170

List of Figures

Figure 1.1 Growth of Enplanements Across the National Airspace System (NAS).....	2
Figure 1.2 Structure of the Main Problem.	6
Figure 2.1 Flowchart of the Methodology Used in Ghobrial and Kanafani (21).....	15
Figure 2.2 Flowchart of the Methodology Used in Hsiao and Hansen (2003).....	17
Figure 2.3 Flowchart of the Passenger Decision Making Process in Market <i>m</i>	36
Figure 3.1 Cost per Available Seat-mile for United and Southwest Airlines (cents per ASM) (Form 10-K, Securities and Exchange Commission).	80
Figure 3.2 Flowchart of the Demand driven Airline and Airport Evolution Study.	81
Figure 3.3 Sample Network from Roanoke, VA (ROA) to San Francisco, CA (SFO). ...	83
Figure 3.4 Sample Network from Roanoke, VA (ROA) to Lewistown, MT (LWT).	84
Figure 3.5 Variation of the Penalty Function with the Volume over Capacity Ratio.	88
Figure 3.6 The Air Traffic Flow Problem.	89
Figure 3.7 Sample Airline Network used for Revenue Computation.	94
Figure 4.1 The Model Integration Process with TSAM.	98
Figure 4.2 The Schedule Building Module.....	101
Figure 4.3 The Schedule Building Module.....	102
Figure 4.4 Calibration Procedure in the O-T-D Analysis.....	107
Figure 4.5 Sample Computation of Revenue between Roanoke,VA(ROA) and Atlanta,GA(ATL).....	126
Figure 5.1 Observed vs. Estimated Route Flows for Trips Originating and Terminating at the Large Hubs.	134
Figure 5.2 Observed vs. Estimated Route Flows for Trips Originating at the Large Hubs and Terminating at the Medium Hubs.....	135
Figure 5.3 Observed vs. Estimated Route Flows for Trips Originating and Terminating at the Medium Hubs.	136
Figure 5.4 Observed vs. Estimated Route Flows for trips Originating and Terminating at the Small Hubs.	137
Figure 5.5 Observed and Estimated Segment Flows for the year 2000.	138
Figure 5.6 Observed Enplanements from the DB1B vs. Estimated Enplanements from the Model Without the Incorporation of Dummy Variables for the Year 2000.	139
Figure 5.7 Observed Enplanements from the DB1B vs. Estimated Enplanements from the Model With the Incorporation of Dummy Variables for the Year 2000.	139
Figure 5.8 Observed and Estimated Segment Flows for the year 2006.	140
Figure 5.9 Observed Enplanements from the DB1B vs. Estimated Enplanements from the Model Without the Incorporation of Dummy Variables for the Year 2006.	141
Figure 5.10 Calculated Frequencies from Formulation I vs. Actual Operating Schedule from the T-100.	142
Figure 5.11 Calculated Frequencies from Formulation II vs. Actual Operating Schedule from the T-100.	143
Figure 5.12 Calculated Frequencies from Formulation III vs. Actual Operating Schedule from the T-100.	144
Figure 5.13 Calculated Frequencies from Formulation IV vs. Actual Operating Schedule from the T-100.	145

Figure 5.14 Calculated Frequencies from Formulation V vs. Actual Operating Schedule from the T-100.	146
Figure 5.15 Calculated Frequencies from Formulation VI vs. Actual Operating Schedule from the T-100.	147
Figure 5.16 Calculated Frequencies from Formulation VII vs. Actual Operating Schedule from the T-100.	148
Figure 5.17 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 1.	149
Figure 5.18 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 2.	150
Figure 5.19 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 3.	151
Figure 5.20 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 4.	152
Figure 5.21 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 5.	153
Figure 5.22 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 6.	154
Figure 5.23 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 7.	155
Figure 5.24 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 8.	156
Figure 5.25 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 9.	157
Figure 5.26 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 10.	158
Figure 5.27 Variation of the Total Hit Rate and the Hit Rate for Cut Segments with the Threshold Value.	160
Figure 5.28 Variation of Fare per Mile and Distance for Coach Class for Passengers Originating in the Large Hubs.	161
Figure 5.29 Variation of Fare per Mile and Distance for Coach Class for Passengers Originating in the Medium Hubs.	162
Figure 5.30 Average Real vs. Average Predicted Fares for all O-D Pairs with a Great Circle Distance less than 500 Miles for the Years 1993-2006.	163
Figure 5.31 Average Real vs. Average Predicted Fares for all O-D Pairs with a Great Circle Distance more than 2500 Miles for the Years 1993-2006.	164
Figure 5.32 Map of Segments Cut in 2007.	171
Figure 5.33 Map of Segments Potentially Added in 2007.	172
Figure 5.34 Map of Segments Cut in 2008.	173
Figure 5.35 Map of Segments Potentially Added in 2008.	174
Figure 5.36 Map of Segments Cut in 2009.	175
Figure 5.37 Map of Segments Potentially Added in 2009.	176
Figure 5.38 Map of Segments Cut in 2010.	177
Figure 5.39 Map of Segments Potentially Added in 2010.	178

Chapter 1 Introduction

1.1) Motivation and Purpose

The air transportation system is an essential part of the transportation infrastructure of the United States. One of the modes that are planned for the future is the Very Light Jet (VLJ) system. Virginia Tech and the National Aerospace and Space Administration (NASA) have developed the Transportation System Analysis Model (TSAM) that predicts the demand for the VLJ mode from the baseline year (2000) to the horizon year (2030). Since TSAM employs a logit model, the demand for VLJ also depends on the utility of the other modes (Automobile, Commercial Aviation, and Rail). Since the only viable competition to SATS over long distances is Commercial Aviation (CA), predicting the evolution of the CA network from the baseline year to the horizon year is of extreme importance.

The growth of air transportation has been rapid across the United States, especially after deregulation in 1978. This growth has in turn placed increasing demand on the transportation system which is quickly becoming saturated. This will lead to increasing delays and lower level of service across the National Airspace System (NAS). Therefore the problem of optimizing the usage of existing infrastructure has assumed greater importance. This cannot be done unless one knows the level of usage of the system, the points of congestion and future trends in the system. To derive these trends, it is necessary to determine the behavior of the system for future years. The growth of air passenger travel in the US is shown in Figure 1.1

The traditional full-service network carriers are facing intense competitive assault on their markets shares by the low-cost airlines, like Southwest. According to the Bureau of Transportation Statistics (2003) Southwest airlines has become the largest airline in the US in terms of passengers carried. In response to their eroding profit and market share, the full-service carriers have taken some drastic measures to cut their operating costs. For example US Airways have ceased operations at Pittsburgh International Airport. And United Airlines has cut frequency from Roanoke, VA to Chicago O'Hare from five flights a day to three flights a day. The scaling back of service by the mainline carriers has served the expansion of Southwest, as Southwest is now the dominant carrier in Pittsburgh. The, modeling the interaction between the low-cost and the traditional carriers is important for understanding the evolution of the aviation industry.

The tragic and catastrophic events of September 11 have caused an appreciable decline in air travel demand for the years 2002-2005. Although it is not possible to predict such tragedies, it is necessary to make an educated estimate about the behavior of the NAS after such catastrophic events.

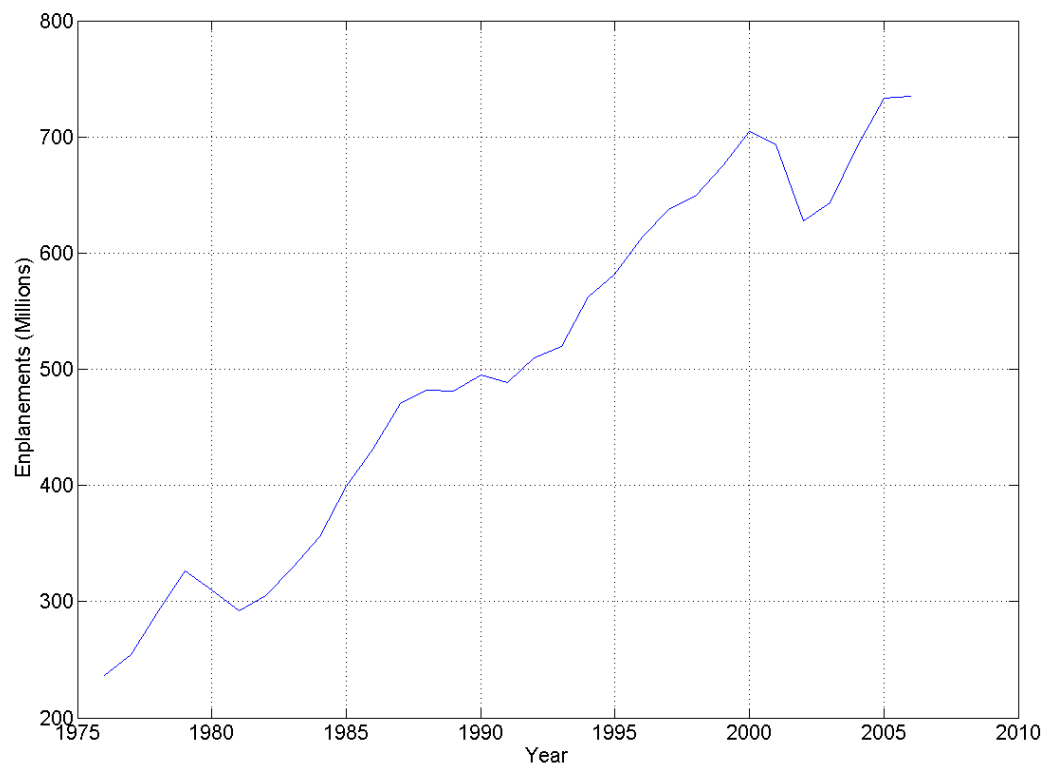


Figure 1.1 Growth of Enplanements Across the National Airspace System (NAS).

1.2) Problem Description

This dissertation focuses on predicting the schedules, fares and network structures of the airline industry every year from 2000 to 2030. The dissertation will treat low cost and full service carriers as single entities as their cost structures have been converging to the same value. The dissertation will also model business and leisure travelers separately since their perceived utility for travel time, fares and other service attributes is different. The model should obtain the probability of travelers choosing a route, for each market, since the demand for each link in the air network is a function of the set of markets connected by the link, and the service attributes of the link. The model should also be able to derive the schedule for each link, since the schedule offered by a carrier is a very important part of the traveler's perceived quality of service. Finally, the model should have the ability to add new links and nodes to the network, when the present network reaches capacity. In short this dissertation can be divided into three sub-problems.

- (1) Air Traffic Network Flow Problem
- (2) Schedule Generator Problem
- (3) Route Generator Problem

1.2.1) Air Traffic Network Flow Problem

The Air Traffic Flow Problem can be stated formally as:

Given a set of Origin-Destination pairs (Airports), the demand between them, the set of paths covering each O-D pair and their service attributes (travel time, fares, frequency, seats) and the set of flight legs in the network

Find the demand for each flight leg, subject to link and node capacity and other additional constraints

The problem of Air Traffic Flow is similar to the Traffic Assignment problem in traffic engineering and transportation planning. The probability of a route being chosen is a function of the service attributes of the route and the capacity offered on the route. For example consider k passengers desiring to travel from i (Origin Airport) to j (Destination Airport) and assume the set R is the set of routes covering the market i - j and D is the demand for the market i - j . Then, the Air Traffic Network Flow Problem can be formulated as

Max

$$Z = f(\text{Passengerutility}(P_R) + \text{Airlineprofit}(P_R)) \quad (1)$$

subject to:

$$\sum_i \sum_j D_{ij} P_R \leq C_R \quad (2)$$

$$\sum_i D_{ij} P_R \leq C_j, \forall j \quad (3)$$

$$\sum_R P_R = 1, \forall i, j \quad (4)$$

1.2.2) Schedule Generator Problem

The schedule generator problem can be stated formally as

Given a set of Origin-Destination pairs (Airports), the arcs covering each O-D pair, the demand for each arc, the fare paid by the passengers in the market, the fleet type and the operating cost for operating an aircraft in the fleet,

Find the schedule set that minimizes the airline operating cost, and satisfies demand, capacity and fleet production constraints

The output of the Air Traffic Network Flow problem (demand for each arc in the network), is the input for the Schedule Generator Problem. Once the optimal schedule has been derived, it may be necessary to have a feedback to the Air Traffic Network Flow Problem, since the attractiveness of a path and hence the probability of the path being chosen depends on the service frequency. For example let $\{\text{Schedules}\}$ represent the schedule set, $\text{Cost}(\{\text{Schedule}\})$ be the cost of operating the schedule, and $\text{Capacity}(\{\text{Schedule}\})$ be the seat capacity of the schedule and $\text{Fleet}(\{\text{Schedule}\})$ be the number of aircraft needed to operate the schedule and Fleetsize be the size of the existing fleet. Then the Schedule Generator Problem can be formulated as

Min

$$Z = \text{Cost}(\{\text{Schedule}\})$$

subject to:

$$\text{Capacity}(\{\text{Schedule}\}) \geq \text{Demand} \quad (6)$$

$$\text{Fleet}(\{\text{Schedule}\}) \geq \text{Fleetsize} \quad (7)$$

1.2.3) Route Generator Problem

The Route Generator Problem can be stated formally as:

Given a base network, the demand for each market in the network, and the fleet type serving each arc in the network, and the direct and indirect costs of building or establishing a new airport

Find the set of routes to be added to the network that minimizes the congestion for the passengers and maximizes the airline profit

Find the set of routes to be removed from the network, which maximizes the increase in profit and minimizes the decrease in revenue

The Route Generator is executed whenever the network flow exceeds the network capacity, congestion at one or more airports becomes excessive or when a new low cost airline enters the market. The output of the Air Traffic Network Flow Model is the input to the Route Generator Model. As in the case of the Schedule Generator the output of the Route Generator Model may have a feedback into the Air Traffic Network Flow Model since the arc and path flow depend on the number of routes for each market. The interaction of three sub-problems is shown in Figure 1.2

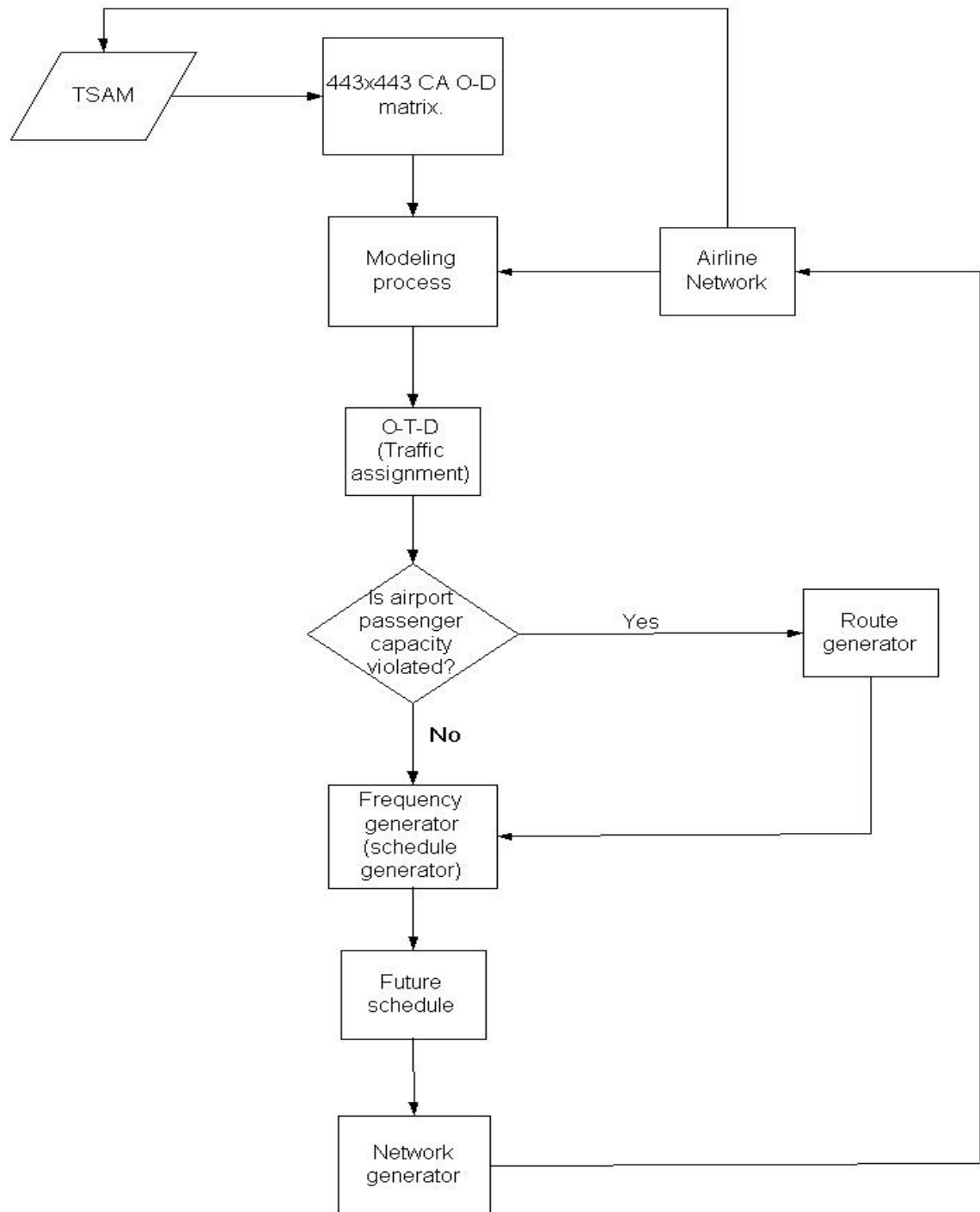


Figure 1.2 Structure of the Main Problem.

1.3) Organization of this Document

This document is organized into three chapters. Chapter 1 discusses the purpose, scope and a brief description of the problem. Chapter 2 discusses the relevant literature and methods pertinent to the Air Traffic Network Flow Problem, Schedule Generator and the Route Generator Problem and a brief discussion of the advantages and drawbacks of each method. Chapter 3 discusses the methodology adopted to solve all three problems

1.4) Contributions of this Study

This study predicts the nationwide trends in Air Transportation System, which includes the airlines, airports and passengers. Specifically, the study can derive trends in fares, demand and the number of operations at each airport. The study can also predict the evolution of new airports and closure of an existing airport. The study also enables the modeler to evaluate the impact of new technologies in aviation: this includes the Very Light Jet (VLJ), tiltrotor and supersonic aircraft. Previous models had only focused on a small subset of the Air Transportation System and did not study the evolution of the entire system. However, this study does not take into account the microscopic features of the system such as airline competition, bankruptcies and mergers. This study also assumes that the aggregate fare trends observed nationwide would apply to individual markets, as deriving separate trends for each market is not practical.

Chapter 2 Literature Review

The literature review section is divided into three parts. In Section 2.1 models that deal with the problem of air traffic network flow are discussed. Section 2.2 outlines models that address the problem of airline schedule design or frequency determination. Section 2.3 deals with studies on choosing the most profitable route or hub location.

2.1) Literature review on the Air Traffic Network Flow Problem

The objective of the Air Traffic Network Flow problem is to determine the number of passengers (demand) for each flight leg (link) and each airport (node) in the network. The problem is particularly difficult since each link in the network is shared between different markets, and therefore carries a heterogeneous mix of passengers. Each set of passengers has its own utility characteristics and consequently differing objectives. The interaction of supply and demand side variables also have to be captured, since airlines decide their schedules and fares based on the projected demand, and passengers choose their carriers and routes based on the supply side variables. The Air Traffic Network Flow Problem can be formally defined as:

Given a set of Origin-Destination pairs, the demand between them, the set of paths covering each O-D pair and their service attributes (travel time, fares, frequency, seats) and the set of flight legs in the network,

Find the demand for each flight leg, subject to capacity and other additional constraints

The Air Traffic Network Flow problem is analogous to the route choice and traffic assignment problems in traffic engineering, and the solutions therefore closely resemble those proposed by ground transportation analysts. However there are important distinctions between ground and air networks.

- 1) The ground networks focus on minimizing the cost or maximizing the benefit to the user. However the Air Traffic Network Flow model has to take into account the interests of both the passenger and the service provider.
- 2) Air network topologies can change quickly relative to traffic networks.
- 3) Competition between various carriers affects the equilibrium solution in air networks, unlike in traffic networks where the effect of competition is minimal.
- 4) Congestion in air networks accumulates at the nodes, on the other hand congestion in traffic networks is concentrated on the links.

However, given these differences the methodologies used to route passengers along ground networks and air networks are quite similar and this section outlines the models used for both the problems. The solution methodologies can be broadly grouped into eight categories

- 1) Entropy based formulations
- 2) Stochastic choice models (Logit models)
- 3) Supply-demand based equilibrium formulations
- 4) Capacity constrained stochastic choice models
- 5) Game theoretic formulations
- 6) Fuzzy logic based formulations
- 7) Heuristic based formulations and
- 8) Probability based formulations.

2.1.1) Entropy Based Formulations

Entropy based formulations are used in transportation planning to determine the trip interchange matrix (no of trips from each origin to destination), given the traffic counts on each link. One of the first studies in Entropy was done by Van Zuylen and Willumsen (1979). The model formulation is:

Let

p_{ij}^a = proportion of trips from i to j, that pass through link a where $0 \leq p_{ij}^a \leq 1$

The fundamental equation of link flow is:

$$V_a = \sum_i \sum_j p_{ij}^a T_{ij} \quad (8)$$

The information contained in a set of N observations, where the state k has been n_k times :

$$I = \log(N!) \pi_k \frac{q_k n_k}{n_k!} \quad (9)$$

Where q_k is the a-priori probability of observing state k.

$$n_{ij}^a = T_{ij} p_{ij}^a \quad (10)$$

Then the a'priori probability of observing state ij on a link a is:

$$q_{ij}^a = \frac{t_{ij} p_{ij}^a}{\sum t_{ij} p_{ij}^a} \quad (11)$$

$$V_a = \frac{\log(V_a!) \prod_{ij} \left(\frac{t_{ij} p_{ij}^a}{S_a} \right)^{T_{ij} p_{ij}^a}}{\prod_{ij} t_{ij} p_{ij}^a} \quad (12)$$

Where $S_a = \sum_i \sum_j p_{ij}^a T_{ij}$

Using Stirling's approximation we get,

$$I_a = \sum_a \sum_{ij} T_{ij} p_{ij}^a \log \left(\frac{T_{ij} S_a}{V_a t_{ij}} \right) \quad (13)$$

subject to:

$$V_a = \sum_i \sum_j p_{ij}^a T_{ij} \quad (14)$$

More recently, Van Aerde and Rakha (2003) and (2005) proposed various flavors of the objective function and evaluated the benefits and shortcomings of each approach. The formulations considered by them are

Max

$$Z(T_{ij}, t_{ij}) = \frac{T!}{\prod_{ij} (T_{ij}!)} \prod_{ij} \left(\frac{t_{ij}}{\sum t_{ij}} \right)^{T_{ij}} \quad (15)$$

Max

$$Z(T_{ij}, t_{ij}) = T \ln \left(\frac{T}{t} \right) - T - \sum_{ij} \left(T_{ij} \ln \left(\frac{T_{ij}}{t_{ij}} \right) - T_{ij} \right) \quad (16)$$

Max

$$Z(T_{ij}, t_{ij}) = - \sum_{ij} \left(T_{ij} \ln \left(\frac{T_{ij}}{t_{ij}} \right) - T_{ij} \right) \quad (17)$$

Max

$$Z(T_{ij}, t_{ij}) = - \sum_{ij} \left(T_{ij} \ln \left(\frac{T_{ij}}{t_{ij}} \right) - T_{ij} + t_{ij} \right) \quad (18)$$

Min

$$Z(T_{ij}, t_{ij}) = \sum_{ij} \left(\frac{1}{2T_{ij}} (T_{ij} - t_{ij})^2 \right) \quad (19)$$

Min

$$Z(T_{ij}, t_{ij}) = \sum_{ij} \left(\frac{1}{2t_{ij}} (T_{ij} - t_{ij})^2 \right) \quad (20)$$

Max

$$Z(T_{ij}, t_{ij}) = \prod_a \frac{V_a!}{\prod_{ij} T_{ij} p_{ij}^a} \prod_{ij} \frac{t_{ij} p_{ij}^a}{V_a} \quad (21)$$

Min

$$Z(T_{ij}, t_{ij}) = \sum_a \sum_{ij} T_{ij} p_{ij}^a \ln \left(\frac{T_{ij} v_a}{V_a t_{ij}} \right) \quad (22)$$

Max

$$Z = T \ln \left(\frac{T}{t} \right) - \sum_{ij} T_{ij} \ln \left(\frac{T_{ij}}{t_{ij}} \right) - \sum_{ij} \left(\lambda_{ij} \cdot 2 \left(\sum_a (V_{ij} p_{ij}^a) - \left(\sum_a p_{ij}^a \left(\sum_{xy} T_{xy} p_{xy}^a \right) \right) \right) \right) \quad (23)$$

The authors point out that, formulations (15) and (21) do not incorporate any simplifying assumptions, while (16), (22) and (23) incorporate only Stirling's approximation. Formulations (17), (18) and (19) incorporate the assumption of constant total demand ($T \approx t$). Formulation (20) incorporates ($T_{ij} \approx t_{ij}$) in addition to the assumptions of (17), (18) and (19). The authors conclude that all the formulations except (17), (18) and (19) produce solutions that correspond to the seed (guess) matrix. The formulations (15), (16), (21), (22) and (23) are independent of the scaling of the seed matrix. Equations (19), (20) are independent of the scaling of the seed matrix if the total number of trips is a constant. Equations (17), (18) are dependent on the scaling of the seed matrix. In the case where the seed matrix is infeasible and the total demand is constant, the formulations (15), (16) and (21) produce optimal solutions. The equations (17), (18), (22) and (23) produce results that are close to the optimal solutions. Equations (19) and (20) are limited in terms of their capability. In the case of infeasible seed matrix and variable total demand the formulations (15) and (16) produce optimal solutions. The equations (17), (21), (22) and (23) produce results that are close to the optimal solutions. Equations (18), (19) and (20) are limited in terms of their capability.

The authors solve a large scale network (1000 zones and 5000 links) in a relatively short time (1 CPU-hour) and conclude that the results are fairly accurate and close to the seed (guess) matrix.

Heidemann (2000) questions the principle of using Entropy to model transportation problems, The author claims that entropy maximization is not suited to the transportation O-D problem since information deficiency occurs from the lack of information on the part of the modeler and not the traveler. The traveler has no entropy or randomness since the traveler knows where to go. Therefore, the study concludes from the traveler point of view there is no entropy.

However the entropy method suffers from the drawback that it does not have any implicit causality factor. In other words, if the travel time or fare on a route changes, the entropy model would be unable to predict the change in the passenger demand for that route.

2.1.2) Stochastic Choice Based Formulations

Logit models have been used extensively to model traveler choice behavior. Logit models have been widely used in traffic engineering and transportation planning to model mode and route choice. However the use of logit models to the Air Traffic Network Flow problem is fairly recent. Weidner (1995) uses a nested logit model to solve the airline hub location problem and predict route and link flows. Weidner assumes that airline hub location decisions are due to econometric reasons alone. The formulation employed by Weidner is:

$$P_m(\text{city } k \mid \text{connecting route}) = \frac{e^{V_{m,k}}}{\sum e^{V_{m,k}}} \quad (24)$$

$$P_m(\text{connect}) = \frac{e^{V_{m,\text{connect}}}}{(e^{V_{m,\text{direct}}} + e^{V_{m,\text{connect}}})} \quad (25)$$

$$V_{m,\text{connect}} = \theta \times I_m \quad (26)$$

$$I_m = \ln \left(\sum e^{V_{m,k}} \right) \quad (27)$$

$$V_m = f(\text{Dist IT}, \text{MaxSegTF}, \text{MinSegTF}, \text{ConPCT}, \text{EPS}, \text{Pax/ASV}, \text{Ops/ASV}) \quad (28)$$

Where,

DistIT: Itinerary Distance.

MaxSegTF : Traffic on maximum segment.

MinSegTF: Traffic on minimum segment.

ConPCT: Connecting passengers at the hub airport.

EPS: Enplanements at the hub airport.

Pax/ASV: No of passengers/capacity.

Ops/ASV: No of operations/capacity.

θ : Inclusive value for the nest.

$P_m(\text{city } k | \text{connecting route})$: Probability of connecting through city k , given that passenger choice is a connecting flight.

$P_m(\text{connect})$: Probability of a passenger choosing a connecting route.

Wei and Hansen (2005) attempt to validate the hypothesis that frequency is more important than aircraft size, in determining the attractiveness of a route. The model used by authors is:

V_{im} : utility of taking up a trip through airline i , in market m

S_{im} : share of an airline i in market m

$$Q_m = D_m \times \left(\frac{\left(\sum_j \exp(V_{jm}) \right)^\theta}{\left(\sum_j \exp(V_{jm}) \right)^\theta + \exp(V_{om})} \right) \quad (29)$$

Where

Q_m : Total trips in market m by air.

D_m : Saturated demand.

The authors assume that

$$\left(\sum_j \exp(V_{jm}) \right)^\theta \ll \exp(V_{om}) \quad (30)$$

Then,

$$Q_m = D_m \times \left(\frac{\left(\sum_j \exp(V_{jm}) \right)^\theta}{\exp(V_{om})} \right) = P_m \times \left(\sum_j \exp(V_{jm}) \right)^\theta \quad (31)$$

$$P_m = \frac{D_m}{\exp(V_{om})} \Rightarrow P_m = \exp(K_m)(Incom_m)^p \quad (32)$$

$$Q_m = \exp(K_m)(Incom_m)^p (Util)^\theta \quad (33)$$

The authors assume the following form for the utility function

$$V_{jm} = a \times \ln(Freq_{jm}) + b \times \ln(Size_{jm}) + c \times \ln(Aval_{jm}) + d \times \ln(Fare_{jm}) \quad (34)$$

The authors then estimate the coefficients by a standard least-squares procedure

The authors conclude that frequency is more important than aircraft size, even when airport congestion is a factor.

Vovsha and Bekhor (1998) suggest the use of a cross-nested logit model for modeling route choice, when a link is shared by more than one route to avoid bias in the estimation.

Another recent development is the issue of aggregated logit models. Unlike disaggregate logit models, aggregate logit models attempt to explain the market share of a particular alternative rather than the individual choice of each person. The aggregate logit model has been applied in market survey analysis in consumer choice models (Chintagunta (2000)). However the application of an aggregate model to air travel passenger choice modeling is a fairly new concept. This scheme was explored by Coldren and Koppelman (2003,2005). However the question of bias introduced due to aggregation of respondent choices is a question that has not been settled yet (Allenby and Rossi (1991)).

The stochastic choice models are better suited to model the route choice problem than the entropy approach, since they employ service attributes to determine the attractiveness of the route. However the major drawback of this approach is that, there is no supply-demand interaction. All the models described above only study the demand side assuming infinite or constant supply. However airlines base their schedules, routes and fares from the projected demand. Therefore a supply-demand model is better equipped to deal with the intricacies of the Air Traffic Network Flow problem.

2.1.3) Supply-demand Based Equilibrium Formulations

Supply-demand models use stochastic choice models to evaluate the demand for each route and link. These models are usually composed of three parts: (a) a discrete choice travelers' choice model for predicting demand, (b) optimization or regression models to decide airlines behavior (c) interaction rules between supply and demand models.

Ghobrial and Kanafani (1995) consider a supply-demand equilibrium model to study the evolution of the airline network under congestion. The authors suggest that by hubbing at a few airports, airlines maintain high levels of aircraft utilization and take advantage of the economies of aircraft size. Passengers also benefit in the form of increased frequency of service. However, increased aircraft operations at major hubs imply certain diseconomies that include congestion delay, increased workload on air-traffic controllers, noise, and pollution. Using a network equilibrium model, the authors attempt to project the future structure of domestic networks and to discuss some policy implications. The authors conclude that network hubbing will continue to persist as an important feature of air transportation, but in a multihub system. Unless expanded substantially, major hubs will suffer from escalated levels of delays and will become mostly traffic-generator airports instead of connecting and transfer points.

The model assumes that the airline network structure depends on

- 1) Increased congestion delays at major hubs.
- 2) New aircraft technologies
- 3) Implementation of hub pricing or slot allocation schemes by the FAA to alleviate congestion.
- 4) Traffic growth between different cities
- 5) Developments of way ports

The effects of the factors (1) to (5) is summarized by the authors as

- 1) Growth in air travel between large cities and other cities will result in more operations at large hubs and more delays
- 2) Growth in air traffic between medium-medium hubs, small-medium hubs and small-small hubs may lead to direct flights between these airports, thus decreasing the demand on large hubs and relieving the congestion. The model assumes that airline markets follow the pattern of monopolistic markets.

The passenger route choice is modeled using a multinomial logit model.

$$T_r = \frac{e^{U(r)}}{\sum_k e^{U(k)}} \quad (35)$$

$U(k) = f(\text{traveltime, fare, frequency, seats, no of stops})$

On account of the interdependence between passenger share and flight frequency on a route, the model assigns flight frequency to the network so that demand and supply are in equilibrium.

The algorithm used in the model is: a) a primary network is first defined by a set of candidate routes for each O-D pair, b) a unit value of frequency is assigned to all routes

and the model calibrated, c) link flows are calculated from route flows, d) frequency is determined to satisfy the load factor e) If the final frequency is not equal to the initial assumed frequency, the initial frequency is set to the final frequency and the process repeated. The model can be represented as a flowchart in Figure 2.1

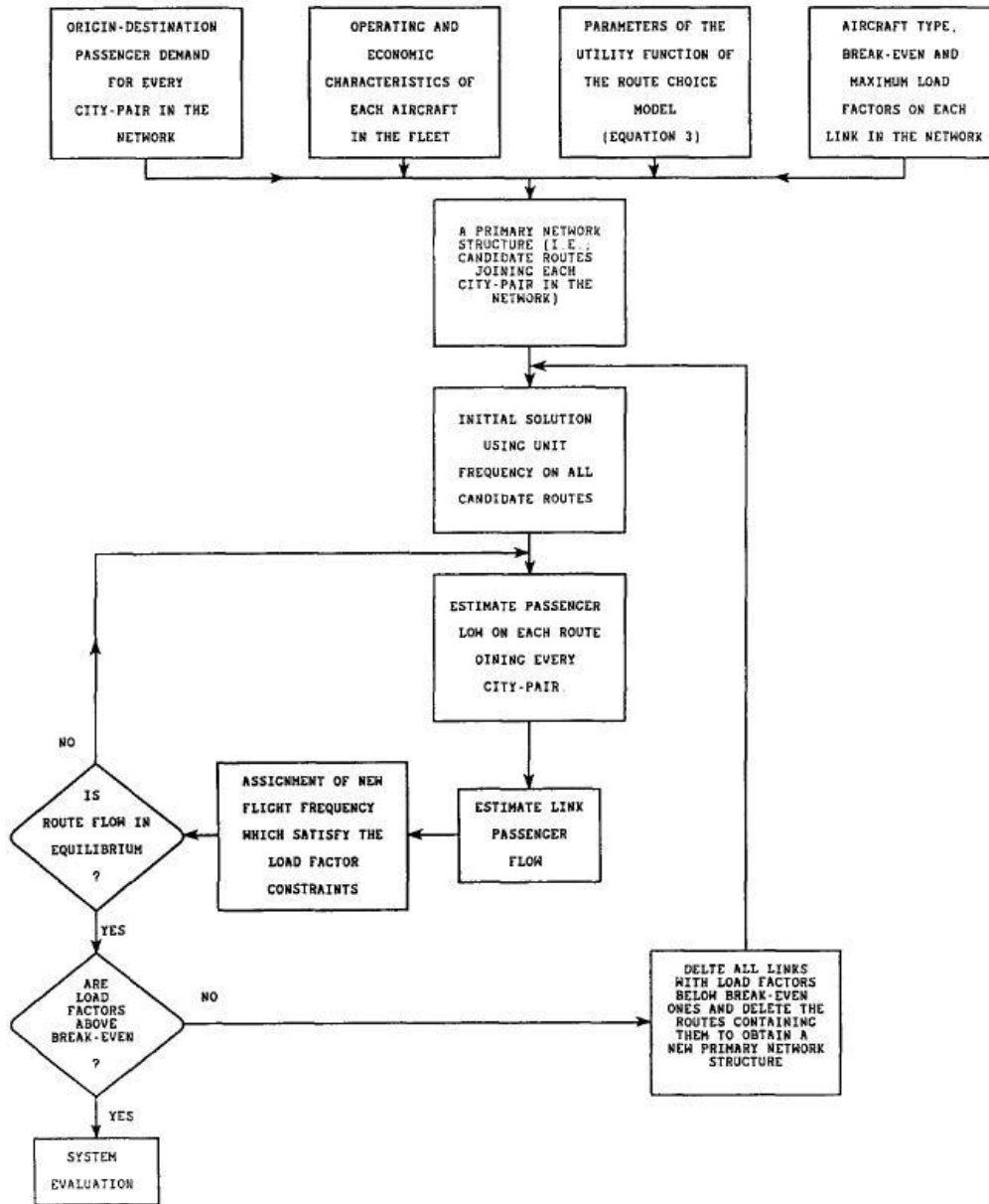


Figure 2.1 Flowchart of the Methodology Used in Ghobrial and Kanafani (21).

Hsiao and Hansen (2003) use an equilibrium model, based on congestion to predict link flows between top 31 airports in the continental US.

Using the DB1B database the authors determine that 66 % of the passengers take direct flights, 32% take one-stop flights and 2% take flights that have 2 or more stops. The authors also determine that 32% of the routes are one-stop, 65% of the routes are more than one-stop and 3% of the routes are non-stop.

The authors point out that there are three varieties of models dealing with passenger flows.

- 1) Supply side only (Mathematical programming approach)
- 2) Demand based models (Logit models).
- 3) Equilibrium models (supply-demand models).

The authors use a nested logit model to predict segment flows from market flows.

$$P_{od}(\text{direct}) = \frac{\exp(V_{od})}{(\exp(V_{od}) + (\sum \exp(\beta * V_{od,i}))^{1/\beta})} \quad (36)$$

$$P_{od,i} = \frac{\exp(V_{od,i})}{\sum \exp(V_{od,i})} \quad (37)$$

$$V_{od} = C_0 + b_{01} \times (D_{od}) + b_{02} \times \ln(Pax_{od}) + b_{03} \times HHI_{od} \quad (38)$$

$$V_{od,i} = C_1 \times Dist_{o-i-d} + C_2 \times \ln(\max pax_{oi/di}) + C_3 \times \ln(\min pax_{oi/di}) + C_4 \times Delay_i \quad (39)$$

The authors model delay using a linear regression model.

$$\ln(Delay_{it}) = d_0 + \sum \alpha_1 \times C_1 + \beta_1 \times \ln(pax_{it}) + \varepsilon_{it} \quad (40)$$

Where:

D_{od} : Direct distance (1000's of miles) between origin and destination.

$Dist_{o-i-d}$: Distance between origin and destination when flow is routed through airport i .

Pax_{od} : Number of passengers on the direct passengers.

$\max pax_{oi/di}$: Number of passengers on the segment with maximum traffic.

$\min pax_{oi/di}$: Number of passengers of the segment with minimum traffic.

HHI_{od} : Herfindahl index of the direct segment.

$Delay_i$: Delay of connecting airport i .

The authors propose the following hypotheses

- 1) The more passengers on a segment, better services such as more frequency will be added by the airlines.
- 2) The longer the distance of the alternative, the less possibility the alternative will be chosen because the longer distance results in greater travel time and increased supply cost.
- 3) The higher the Herfindahl index means that less competition exists in this segment. Airlines may have lesser incentive to provide better service and charge lower fares on this segment.
- 4) The greater the airport delay, the lesser chance of the airport being chosen for connection.

The authors use all non-stop and one-stop routes for the 31 benchmark airports. The algorithm used is: a) initially it is assumed that all passengers take direct flights and segment flows calculated (SegPax^0), b) the route choice model is then used to calculate the new values of the segment flows (SegPax^1), c) then SegPax^1 is used as the initial value to get the next value of the segment flows (SegPax^2) and d) Process is executed till $\text{SegPax}^n \approx \text{SegPax}^{n+1}$.

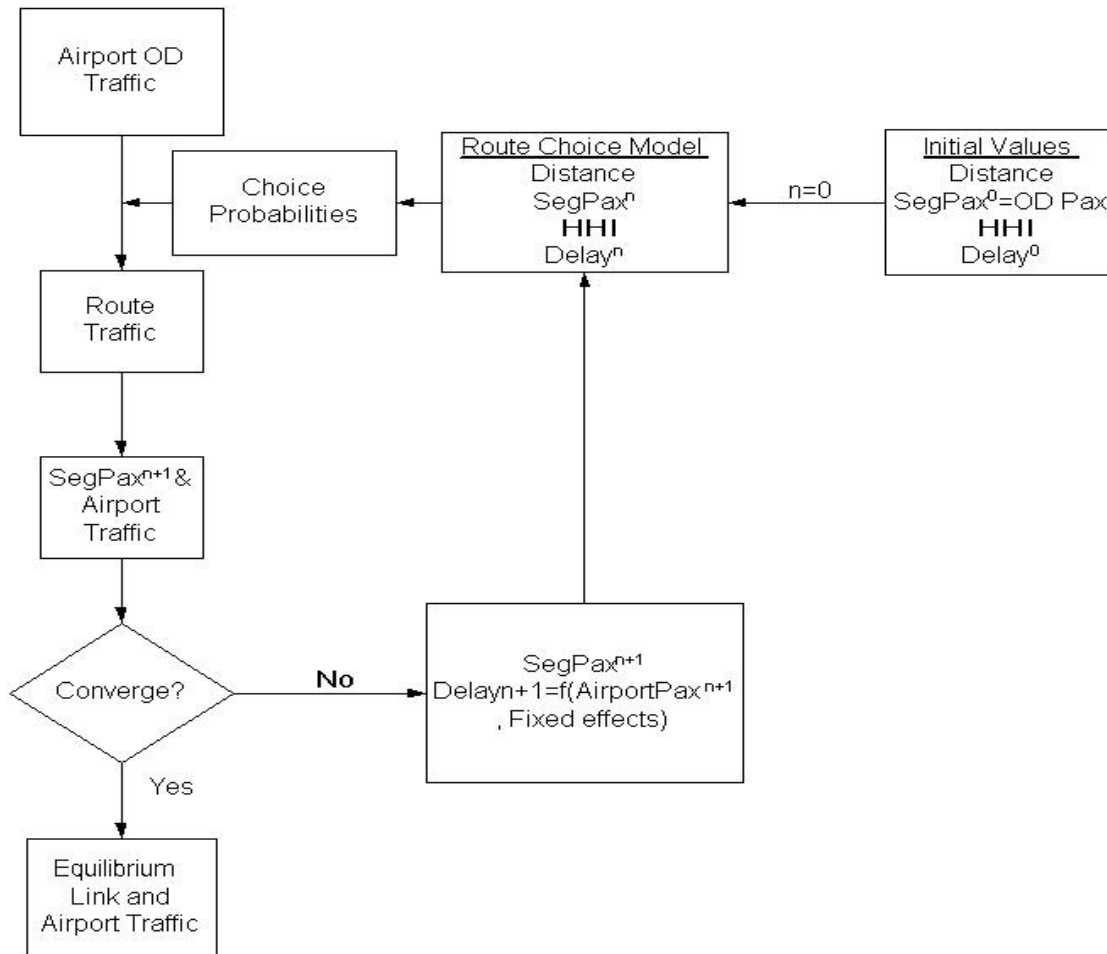


Figure 2.2 Flowchart of the Methodology Used in Hsiao and Hansen (2003).

Hsu and Wen (2005) determine flight frequencies on an airline network with demand–supply interactions between passenger demand and flight frequencies. The model consists of two submodels, a passenger airline flight choice model and an airline flight frequency programming model. The demand–supply interactions relevant to determining flight frequency on an airline’s network are analyzed by integrating these two submodels

The demand model used by the authors is:

- t_{rsp}^x : Average block time spent flying route p on airline x between O-D pairs r and s .
- sd_{rsp}^x : Scheduling delay on flying route p on airline x between O-D pairs r and s .
- N_{rsp}^x : Frequency on flying route p on airline x between O-D pairs r and s .

The total cost is:

$$\tau_{rs} (t_{rsp}^x + v^{sd} sd_{rsp}^x) \quad (41)$$

Where:

- τ_{rs} : Value of time.
- v_{sd} : Multiplier for schedule delay.
- ϕ_{rs}^x ($0 \leq \phi_{rs}^x \leq 1$) is the random variable denoting the airfare discount factor from frequent flyer programs.
- tp_{rsp}^x : Basic airfare.

The generalized cost is:

$$G_{rsp}^x = (1 - \phi_{rs}^x) tp_{rsp}^x + \tau_{rs} (t_{rsp}^x + v^{sd} sd_{rsp}^x) \quad (42)$$

The authors assume a correlation between ϕ_{rs}^x and τ_{rs} , and model this using a joint pdf between the two variables ($f(\phi_{rs}^x, \tau_{rs})$).

Let there be i groups of passengers with the same value-of-time τ_{rsi} and j groups with similar frequent flyer programs ϕ_{rsj}^x .

For every group (i, j) the generalized cost is

$$G_{rsp(i, j)}^x = (1 - \phi_{rsj}^x) tp_{rsp}^x + \tau_{rsj} (t_{rsp}^x + v^{sd} sd_{rsp}^x) \quad (43)$$

The probability of choosing airline 0 for travel between r - s using route p is

$$\Pr(G_{rsp}^0 \leq G_{rsp}^x) = \sum \sum \Pr(G_{rsp}^0 \leq \min(G_{rsp(i, j)}^x)) p(i, j) \quad (44)$$

For the multiple routes by the same airline between same O-D pair, a route is chosen only if its generalized cost is less than the generalized cost of the all routes.

$$MS_{rsp}^0 = \Pr\{G_{rsp}^0 \leq G_{rsk}^0\} \times \Pr\{G_{rsp}^0 \leq G_{rsp(i,j)}^x\} \quad (45)$$

The expression for the above equation is given in the attached image in (2)

The total number of f_{rsp}^0 carried by the object airline on its route p between OD pair r-s can be estimated as

$$f_{rsp}^0 = F_{rs} MS_{rsp}^0 \quad (46)$$

The total number of passengers carried by the object airline between OD pair r-s

$$f_{rs}^0 = \sum_p f_{rsp}^0 \quad (47)$$

The supply model is:

Variables used

$G(N,A)$ is a graph, with N representing the nodes and A representing the set of arcs.

R and S are the set of origins and destinations respectively.

N_{rspq}^0 = frequency of airline 0 (object airline) between r and s for route p for aircraft type q.

$$N_{rsp}^0 = \sum_q N_{rspq}^0 \quad (48)$$

$$N_{rs}^0 = \sum_p \sum_q N_{rspq}^0 \quad (49)$$

Y_{aq}^0 = frequency offered by airline 0 on link a by aircraft type q.

$$Y_{aq}^0 = \sum_{r,s} \sum_p \delta_{a,p,q}^{r,s} N_{rspq}^0 \quad (50)$$

Where:

$\delta_{a,p,q}^{r,s} = 1$ if a link a is part of route p served by aircraft type q between r-s.
 $= 0$ otherwise.

The total flight frequency by airline 0 on link a

$$Y_a^0 = \sum_q Y_{aq}^0 \quad (51)$$

The total link flow through link a for airline 0 is

$$f_a^0 = \sum_{r,s} \sum_p \delta_{a,p,q}^{r,s} f_{rsp}^0 \quad (52)$$

Where:

$\delta_{a,p}^{r,s} = 1$ if link a is a part of route p between r-s.
 $= 0$ otherwise.

The aircraft operating cost is written as

$$C_a^T(Y_{aq}^0) = \sum_q C_{aq}^D(Y_{aq}^0) + C_a^T(Y_{aq}^0) \quad (53)$$

Where:

C_{aq}^D : Direct operating costs for aircraft type q on link a .

C_a^T : Indirect cost.

The frequency programming problem is

Max

$$\pi_0 = \sum_{r,s} \sum_p (1 - E(\phi_{rs}^0)) t p_{rsp}^0 f_{rsp}^0 - \sum_a C_a^T(Y_{aq}^0) \quad (54)$$

subject to:

$$\sum_q n_q l_a Y_{aq}^0 - \sum_{r,s} \sum_p \delta_{a,p,q}^{r,s} f_{rsp}^0 \geq 0 \quad (55)$$

$$Y_{aq}^0 = \sum_{r,s} \sum_p \delta_{a,p,q}^{r,s} N_{rspq}^0, \forall a, q \quad (56)$$

$$f_{rsp}^0 = F_{rs} M S_{rsp}^0, \forall r, s, p \quad (57)$$

$$Y_{aq}^0, N_{rspq}^0 \text{ integer}$$

Constraints (1), (2) and (3) represent capacity constraint, frequency summation constraint and aircraft utilization constraint respectively. The authors apply this model to the network of China Airlines (CI) of Taiwan.

The supply-demand model uses the supply and demand models in an iterative fashion to find the final equilibrium solution. The supply-demand models use both the supply and demand side to derive a solution in which both supply and demand are in equilibrium and therefore, is a fairly accurate representation of the airline network. However the drawback of this method is that the quality of the solution depends on the rate of the convergence of the algorithm, which could be slow for large networks.

2.1.4) Game-Theoretic Formulations

Game-theoretic formulations attempt to achieve the same objective that supply-demand equilibrium frameworks achieve, i.e. game-theoretic methods model both the supply and demand side variables and the interactions between the two components. However game theory based methods unlike supply-demand models attempt to solve for the equilibrium solution in a single step rather than derive the solution in an iterative manner. The game theory methods assume that each airline is a player in a non-cooperative game and solve for the dominant strategy of each player. The strategy of each player is a function of

supply side variables (network structure, fares, frequency and aircraft type) and demand side variables (passenger flow).

Game theory has been used fairly widely in literature to model the competition between firms. However the use of game theory to model airline competition is fairly recent. Hansen (1990) uses a simplified non-cooperative game to model the behavior of the airline network in the year 1985.

The author makes the following assumptions to reduce the computational effort involved:

- 1) Only the airlines are considered as players, no airports, travelers or unions are modeled.
- 2) Airline payoff functions are simplified.
- 3) Hub locations are exogenous.
- 4) Capacity allocated to each fare class is fixed.
- 5) No fleet constraints.
- 6) O-D demand is inelastic to price and service levels.

The variables used in the model are:

np : No. of points in the set.

D_{ij} : Demand between the points (symmetric matrix).

There are $np+1$ classes, each player falls into one of those classes.

Class 0 – Low cost (point-to-point airlines).

Class i – Full service carriers.

Q_{ij} : Demand between airports i and j .

L_{ij} : Variable for the service offered by the competitor in the same market.

The profit function of the Low Cost Carriers (LCC's) are

$$\pi^0 \{v_{ij}\} = \sum_{i=1}^{np} \sum_{j=1}^{np} (S^{ns}(v_{ij}, L_{ij}) \times Q_{ij} \times F_{ij} - C(v_{ij}, S^{ns}(v_{ij}, L_{ij}), Q_{ij})) \quad (58)$$

Where:

v_{ij} : Frequency between i and j .

F_{ij} : Fare in market i and j .

S_{ns} : Market share.

For a hub airline the profit function is given by:

$$\begin{aligned} \pi^k \{v\} = & \sum_{i=1}^{np} \sum_{j=1}^{np} [\delta_{kj} \cdot S^{ns}(v_j, L_{ij}) + \\ & (1 - \delta_{kj}) \cdot S^{hs}(v_j, v_i, D_{ik} + D_{kj} - D_{ij}, L_{ij})] F_{ij} \cdot Q_{ij} - 2 \sum_{i=1}^{np} C(v_i, \sum_{j=1}^{np} [\delta_{kj} \cdot \\ & S^{ns}(v_j, L_{ij}) + (1 - \delta_{kj}) \cdot S^{ns}(v_j, v_i, D_{ik} + D_{kj} - D_{ij})] Q_{ij} \cdot D_{ij}) \end{aligned} \quad (59)$$

Where:

$$\delta_{kj} = 1 \text{ if } k=j \\ = 0 \text{ otherwise}$$

To execute this game, the author has relationship for S^{ns} , S^{hs} and C .

$$S_{ij}^m = \frac{\exp(V_{ij}^m)}{\sum \exp(V_{ij}^m)} \quad (60)$$

For direct service

$$V^d = \alpha F + \varphi_0 \ln(v) + V_{dct},$$

Where

F : Fare.

v : Frequency.

V_{dct} : Consumer preference for direct service.

For hub-service

$$V^{hs} = \alpha F + \varphi_1 \ln(v_{\max}) + \varphi_2 \ln(v_{\min}) + \sigma(CIRC) \quad (61)$$

Where:

F : Fare.

$v_{\max}$: Frequency of the most frequent leg of the itinerary.

$v_{\min}$: Frequency of the least frequent leg of the itinerary.

$CIRC$: Circuitry.

The fare between an O-D pair is expressed as:

$$\frac{F_{ij}}{D_{ij}} = \alpha + \frac{\beta}{D_{ij}} + \frac{\gamma}{D_{ij}^2} \quad (62)$$

Where:

F_{ij} : Fare between the airports i and j .

D_{ij} : Great circle distance between i and j .

The aircraft cost function is assumed to be of the form.

$$C(v, Q, D) = v.C^f(S(Q, R_d), D) \quad (63)$$

Where:

v : Frequency of service

D : Great circle distance

Q : Passenger link flow

S : Number of seats per flight

R_d : Denial rate per flight

The functional form of R_d is

$$R_d = 0.455 \left(\frac{4.12}{\sqrt{Q}} \right)^{0.645} \left(\frac{4.12\sqrt{Q}}{(S-Q)} \right)^{1.79} \quad (64)$$

The number of seats in each flight is calculated as,

$$S(Q, R_d) = Q + R_d \left(\frac{Q^{0.32}}{R_d^{0.56}} \right)$$

The author solved the game-theoretic formulation by maximizing the profit of each airline given the strategy of the other airlines. The quasi-equilibrium state was found to resemble the actual system with regard to a number of key system variables, such as the proportion of passengers using connecting service, and with regard to activity levels at most of the largest hubs. On the other hand, there were substantial divergences with respect to some system variables, and with respect to the levels of activity at hubs of two types: those located in multi-airport regions, and those with comparatively weak local markets. The author applies this model to 205 markets from a sample of 21 cities.

Hong and Parker (1992) develop a model which is aimed at developing an efficient air traffic system for given demand and airport capacity levels by the proper pricing of landing slots. A computable Nash equilibrium model is used in the context of a two-stage, game-theoretic representation of a market mechanism for slot allocation. A variational inequality formulation is then used to solve this oligopolistic air transport market model. The choice of travelers among competing airlines is represented by a logit model.

Two models are proposed for pricing of landing slots: one with an exogenously determined allocation, and the second with an endogenous allocation. Each model derives the flight patterns, ticket prices, routes, and carrier choice for passengers, and landing fees. The authors use a Nash-equilibrium two stage game. The scheduling committee assigns the number of landing slots to each airline at each airport. Then the equilibrium model is used to determine flight schedules, fares and slot prices

Variables used:

i : Origin nodes.

j : Destination nodes.

K : Set of airlines.

N_k : The set of nodes that airline k has under control.

A_k : The set of nodes that airline k has under control.

A : The set of all arcs.

a : A single instance of A .

$G(N, A)$: The air transportation network.

W : The set of O-D pairs.

W_k : The set of O-D pairs served by airline k .
 w : A single instance of W .
 P_w : The set of routes for W .
 $P = \cup P_w$.
 p : A single instance of P .
 Ft^{ka} : Number of flights in the planning horizon flown on a for all k and t .
 Ft^{kp} : Number of flights in the planning horizon flown on p for all k and t .
 $F_k : (Ft^{ka} \text{ or } Ft^{kp})$.
 ε : Maneuvering and ground control time per flight.
 ρ_p^k : Coefficient of schedule efficiency of airline k on route p .
 $\lambda_{pn} = 1$ if route p goes through airport n
 $= 0$ otherwise
 $\lambda_{ap} = 1$ if arc a is used on route p
 $= 0$ otherwise
 T : Set of aircraft types.
 t : An instance of T .
 ALF^k : The average load for aircraft k .
 LD_j : The number of landing rights for airport j .
 TO_i : The number of takeoff rights for airport i .
 ST_p^k : Scheduled travel time on route p for OD pair w .
 ST_a^k : Scheduled travel time on arc a .
 CD_{pn}^k : Connecting delay on node n on route p .
 AT_p : Actual travel time on route p .
 TST_k : Maximum utilization of aircraft type t .
 I_{nt} : Maximum number of aircraft of type t which can be stored at airport n for a single time-period.
 SA_n : Net supply of aircraft at airport n .
 UF_t : Total aircraft of type t .
 R_p^k : Rate charged for a unit of service on route p for OD pair w .

The model formulation is:

The utility function is given by

$$V(p) = \alpha ST_p^k + \beta CD_p^k + \gamma R_p^k + \eta F_p^k + \mu \phi_p^k + \varphi_p^k \quad (65)$$

Where:

$V(p)$ = disutility function.

ϕ_p^k : type of aircraft flown on route p .

φ_p^k : travel pattern of route (direct, one-stop or two-stop).

Probability of choosing route p on airline k

$$P = \frac{e^{-V(p_w^k)}}{\sum \sum e^{-V(q)} + \chi} \quad (66)$$

Where:

χ = outside good.

$D_p^k = D_w * \Pr(p_w^k)$.

Where:

D_p^k = demand on route p , on airline k .

Each airline k has the following constraints

$$1) F_{ta}^k = \sum_{p \in P} \delta_{ap} F_{tp}^k; F_{ta}^k \geq 0, F_{tp}^k \geq 0, \forall t, a, k \quad (67)$$

$$2) \sum_{w \in W} \sum_{p \in P} \delta_{ap} D_p^k \leq \sum_a \sum_t \delta_{ap} (AFL_{ta}^k) S_t F_{ta}^k, \forall k, w, p \quad (68)$$

$$3) AT_p^k = \sum_a \delta_{ap} ST_a^k + \sum_n \lambda_{pn} CD_n, \forall k, w, p \quad (69)$$

$$4) \sum_t \sum_n \sum_a F_{ta}^k \leq LD_j^k, \forall k, j \quad (70)$$

$$5) \sum_t \sum_n \sum_a F_{ta}^k \leq TO_i^k, \forall k, i \quad (71)$$

$$6) \sum_{i \in N^k} \sum_a F_{ta}^k + SA_n \leq \sum_a \sum F_a^k, \forall k, n, t \quad (72)$$

$$7) \sum_{i \in N^k} \sum_{a=(i,n)} F_{ta}^k - \sum_{j \in N^k} \sum_{a=(n,j)} F_a^k = I_{tn}^k, \forall k, n, t \quad (73)$$

$$8) \sum_{a \in A} ST_{ta}^k (1 + \varepsilon_{ta}^k) F_{ta}^k \leq TST_t UF_t^k, \forall t \quad (74)$$

The objective function of each airline becomes

Max

$$\begin{aligned} \pi^k(X^k, X^{-k}) &= REV^k(F^k) - C^k(F^k) - C^k(UF^k) \\ &= \sum_{w \in W} \sum_{p \in P} \Pr(p) \times D_w \times R_p^k - \sum_{a \in A} \sum_{t \in T} F_{ta}^k \times C_{ta}^k(F) - \sum_{t \in T} C_t^k(UF) \times UF_t^k \end{aligned} \quad (75)$$

Where:

$Rev^k(F^k)$ = revenue generated by airline k for flights k

$C^k(F^k)$ = Total cost incurred in flying flights F^k
Where:

- Constraint (1) satisfies arc and path flow definition.
- Constraint (2) satisfies the supply-demand constraint on each arc.
- Constraint (3) satisfies the flight time definition.
- Constraint (4) satisfies the landing rights constraint for airline k at airport j .
- Constraint (5) satisfies the takeoff rights constraint for airline k at airport i .
- Constraint (6) satisfies the flow conservation constraint for airport n .
- Constraint (7) satisfies the aircraft inventory constraint at airport n .
- Constraint (8) satisfies the aircraft availability for aircraft type t .

In the second formulation the total number of slots available at an airport is given, and the slots available to each airline are endogenous to the model. The constraints (4) and (5) now become

$$\sum_{k \in K} \sum_{t \in T} \sum_{n \in N} \sum_{a=(n,j) \in A} F_{ta}^k \leq LD_j, \forall j \quad (76)$$

$$\sum_{k \in K} \sum_{t \in T} \sum_{n \in N} \sum_{a=(i,n) \in A} F_{ta}^k \leq TO_i, \forall i \quad (77)$$

The authors solve the above game for a three airline, three city network and conclude that due to the small nature of the test network, network-wide effects evident in large-scale models would not appear in their results.

Dobson and Lederer (1993) use a combination of supply-demand based equilibrium formulations and game theory to derive the routes and schedules of competing airlines in a hub-and-spoke system. Their method finds the best flights, routes and flights in a hierarchical process. The third level which is the lowest, finds the prices that satisfy capacity constraints and consumer choice behavior. It also determines the airline's total revenue for all the routes. The second level generates a set of feasible routes from the set of flights chosen by the airline. The first level searches over the set of flights to find the profit-maximizing set. The model for profit maximization (level three) is:

Variables used:

- F - The fixed cost of aircraft in dollars per day.
- V - The variable cost in dollars per aircraft hour.
- M - The aircraft capacity.
- w - The value of time for schedule delay.
- v - The value of time for travel time.
- p - The fair paid by the passengers.
- $C_1, C_2, \dots, C_n$: set of cities in the network.

Let

$\beta(C_i, C_j, t)$ = density of customers whose most preferred time of departure from C_i to C_j is time t .

$R(C_i, C_j)$ = Airline's routes between C_i and C_j .

$RC(C_i, C_j)$ = Competitor's routes between C_i and C_j .

Then the demand for route r is:

$$d_r(t) = \beta_k \frac{e^{-\alpha(\omega|t-t_r| + \nu_r + p_r)}}{\sum_{r'=1}^{|R|+|RC|} e^{-\alpha(\omega|t-t_{r'}| + \nu_{r'} + p_{r'})}} \quad (78)$$

The total demand for route r for the entire day

$$d_r = \int_0^{24} d_r(t) dt \quad (79)$$

The authors derive a closed form expression for d_r as:

$$d_r = \sum_{k=1}^H \frac{\beta_k}{2\alpha\omega} \frac{C_{rk}}{CP_k} \log(e^{-2\alpha\omega t} CP_k + CN_k) \Big|_{a_k}^{a_{k+1}} \quad (80)$$

if $r \in P_k$, and

$$d_r = \sum_{k=1}^H \frac{\beta_k}{2\alpha\omega} \frac{C_{rk}}{CN_k} \log(e^{2\alpha\omega t} CN_k + CP_k) \Big|_{a_k}^{a_{k+1}} \quad (81)$$

if $r \in N_k$

Where:

$$P_k = \{r' \in R \cup RC \mid \omega_{r'k} = \omega\}, \text{ and } N_k = \{r' \in R \cup RC \mid \omega_{r'k} = -\omega\}$$

$$CP_k = \sum_{j \in P_k} C_{jk} \quad \text{and} \quad CN_k = \sum_{j \in N_k} C_{jk}$$

The profit maximization routine is

Max

$$\sum \sum p_r d_r - V \sum t_f - FN_p \quad (82)$$

subject to:

- (i) the variable d_r satisfies equations (1) and (2)
- (ii) $A'd \leq Me$

Where:

M is the capacity of the aircraft.

The matrix A is such that
 $a_{rf} \in A = 1$ if flight f is a part of route p .
 $= 0$ otherwise.

The lagrangian of the problem can be written as

$$L(p, \lambda) = \sum \sum p_r d_r - \sum \lambda_f (\sum \sum d_r - M) \quad (83)$$

The authors note that the above problem is a global nonconvex optimization problem, and therefore the authors derive the lagrangian of the above problem and arrive at a solution using sub-gradient optimization.

The level two subroutine consists of deriving viable routes from the flight schedule. The input parameter for the algorithm is the maximum waiting time at the connecting airport and the output is the set of all routes that cover the market satisfying the maximum waiting time constraint.

The level one subroutine consists of choosing the optimal, profit-maximizing set of flights. The authors divide a day into a set of equal time intervals. The authors adopt a different approach to solve the scheduling and fleet assignment problem concurrently. Initially all possible sets of flights between each pair of city is considered. The number of flights needed to cover the initial schedule. For example, if it takes k time-period to fly to the city and back, then k aircraft are needed to cover all flights to and from that city for one day. In graph theory terms the authors start with an Eulerian graph (graph in which an Eulerian cycle can be constructed).

An aircraft's schedule is generated by assigning an aircraft to the node, whenever that node is crossed by the cycle. This produces a periodic aircraft schedule with the period the same as the period of the Eulerian cycle. The initial schedule from the Eulerian cycle is modified by using the following three steps

- 1) A pair of consecutive flights, one flying in to the hub from the city and flying out to the city again, is replaced by a pair of flights, one flying in to the hub from the city and the next flight flying from the hub to itself (aircraft is idle at the hub).
- 2) A pair of consecutive flights, one flying in to the city from the hub and flying out to the hub again, is replaced by a pair of flights, one flying in to the city from the hub and the next flight flying from the city to itself (aircraft is idle at the city).
- 3) Tours with a period of one day is eliminated.

These modifications preserve the Eulerian property of the graph and produce a periodic aircraft schedule.

The authors use a greedy algorithm to choose the set of flights for elimination.

Using the original unaltered schedule, let the optimal price and lagrangian solution be (p^*, λ^*) . Using Rule 1, the model eliminates two flights f and f' . The new optimal solution is given by $p^{f'}, \lambda^*$. The cost-savings from the eliminating the two flights is

$$\Delta_{ff} = L'(p^{f'}, \lambda) + V(t_f + t_{f'}) - L(p, \lambda)$$

The same computation is repeated for all pairs of flights and the pair of flights with the maximum value of Δ_{ff} are eliminated and the model re-run to estimate profits. If profits

have increased then the two flights are eliminated, or else they are incorporated in the schedule. The same elimination procedure is used on flight pairs in the second case.

Oum, Zhang and Zhang (1995) derive a model to examine the reasons for the preponderance of hub-and-spoke networks and their effect on the profitability and market share vis-à-vis linear (point-to-point) networks. They assume three cities and look at a hub-and-spoke and linear network, and solve for the duopoly equilibrium conditions.

The authors claim that there are two advantages of a hub-and-spoke system

- 1) Economies of scale (cost per passenger demand).
- 2) The network enables service to a greater number of OD pairs, with more frequency capturing greater market share.

The disadvantages of a hub-and-spoke network are:

- 1) Dominant carriers in hub-and-spoke networks produce barriers for entry.
- 2) Passengers pay more fare for flights to and from a hub, the so-called "hub-penalty".

The model consists of two parts

- 1) Authors examine the benefits of switching from a linear to a hub-and-spoke network.
- 2) Authors present a framework to analyze the competition between two oligopolistic network oriented firms.

Authors assume three cities (I,H,J Hub=J) and assume two carriers(i=A,B) and two networks

- 1) A hub-and-spoke network with its hub at H
- 2) Point-to-point (linear network)

The model is a two-stage game consisting of

- 1) Carriers choose their network structure.
- 2) After choosing their network structure, carriers choose their output levels.

The model formulation is:

For the linear network

$$Cost = \sum_{k=1}^3 C_k^i(x_k^i) \quad (84)$$

For the hub-and-spoke network

$$Cost = \sum_{k=1}^2 C_k^i(X_k^i) + C_h^i \quad (85)$$

Where $X_k^i = x_k^i + x_3^i$, and $C_h^i = C_d^i + C_h^i$

Demand is given by $x_k^A = D_k^A(\rho_k^A, \rho_k^B)$ and $x_k^B = D_k^B(\rho_k^A, \rho_k^B)$ Where ρ denotes the ticket prices. Inverting the above relation we get,

$\rho_k^A = d_k^A(x_k^A, x_k^B)$ and $\rho_k^B = d_k^B(x_k^A, x_k^B)$

The model assumes,

$$\frac{\partial d_k^i}{\partial x_k^j} < 0 \quad (86)$$

Let $g_k^i(x_k^i)$ = congestion cost in the network.

Cost for a linear network

$$\rho_k^{IL} = p_k^{IL} + g_k^i(x_k^i), k=1, 2, 3 \quad (87)$$

Cost for a hub-and-spoke network

$$\rho_1^{IH} = p_1^{IH} + g_1^i(X_k^1) \quad (88)$$

$$\rho_2^{IH} = p_2^{IH} + g_2^i(X_2^i) \quad (89)$$

$$\rho_3^{IH} = p_3^{IH} + g_2^i(X_2^i) + g_1^i(X_k^1) + \gamma^i \quad (90)$$

The profit for a linear network is

$$\pi^{iL}(x^A, x^B) = \sum_{k=1}^3 d_k^i(x_k^A, x_k^B) x_k^i - \sum_{k=1}^3 C_k^i(x_k^i) - \sum_{k=1}^3 g_k^i(x_k^i) x_k^i \quad (91)$$

The profit for a hub-and-spoke network is

$$\pi^{iH}(x^A, x^B) = \sum_{k=1}^2 d_k^i(x_k^A, x_k^B) x_k^i - \sum_{k=1}^2 C_k^i(X_k^i) - \sum_{k=1}^2 g_k^i(X_k^i) X_k^i - C_h^i - \gamma^i \quad (92)$$

The authors point out that

$$\partial^2 \pi^{iH} / (\partial x_k^i \partial x_3^i) > 0,$$

The above result proves that if local output increases, connecting output decreases and vice-versa.

The authors assume that products of the two firms are substitutes for each other in each city-pair market.

Therefore,

$$\partial^2 \pi^i / (\partial x_k^i \partial x_j^i) < 0.$$

The authors then introduce the following function for ease of computation

$$\pi^i(x^A, x^B, \theta^i) = \theta^i \pi^{iH}(x^A, x^B) + (1 - \theta^i) \pi^{iL}(x^A, x^B) \quad (93)$$

Where:

$$\begin{aligned} \theta^i &= 1 \text{ if the network is a hub-and-spoke network.} \\ &= 0 \text{ otherwise.} \end{aligned}$$

The cost differential between the linear and the hub-and-spoke network can be written as

$$\begin{aligned} C^L - C^H &= \left(\sum_{k=1}^3 C_k^i(x_k^i) + \sum_{k=1}^3 g_k^i(x_k^i) x_k^i \right) - \\ &\left(\sum_{k=1}^2 C_k^i(X_k^i) + \sum_{k=1}^2 g_k^i(X_k^i) X_k^i + C_h^i - \gamma^i \right) \end{aligned} \quad (94)$$

Which can be rewritten as

$$C^L - C^H = \left[\left(\sum_{k=1}^3 C_k(x_k) - \left(\sum_{k=1}^2 C_k(X_k) + C_h \right) \right) + \left[\sum_{k=1}^3 g_k(x_k) - \left(\sum_{k=1}^2 g_k(X_k) X_k + \gamma x_3 \right) \right] \right] \quad (95)$$

The first bracket in the term represents production costs, assuming the production costs are the same the first bracket vanishes. Then the relative profitability of the two networks will depend on the density of the market, if the spoke-spoke market is thin, then hubbing represents a significant increase in service quality for that market in terms of improved flight frequency. On the other hand if the spoke-spoke market is a large market then hubbing causes substantial inconvenience in terms of connecting times.

According to the model, the strategic advantages of hubbing are:

- 1) Increases the market share of the carrier (provided the marginal cost of the hub-and-spoke network is less than that of the linear network)
- 2) Prisoner's dilemma for both players and
- 3) Decreases the profit of the rival

Hendricks, Piccione and Tan (1999) study a model in which two carriers choose networks to connect cities and compete for customers. The authors show that if firms compete aggressively (Bertrand type behavior), one carrier operating a single hub-and-spoke network is an equilibrium outcome. Competing hub-and-spoke networks are not an equilibrium outcome, but duopoly equilibria in non-hub networks can exist. If carriers do not compete aggressively equilibrium with competing hub-and-spoke networks exists as long as the number of cities is not too small. The authors provide conditions in which all equilibria consist of hub-and-spoke networks.

The authors use a two-stage game, to examine when hub-and-spoke networks become conditions. The authors use two carriers

- 1) When carriers compete aggressively after choosing their networks, then the equilibrium is a monopoly where all O-D pairs are served by a single carrier operating a hub-and-spoke network, duopoly hub-and-spoke networks are not possible.
- 2) Duopoly equilibria exist in non-hub networks
- 3) When carriers compete non-aggressively, equilibrium is possible in hub-and-spoke networks

The model formulation is:

The variables used are,

i: (A,B) : The set of carriers.

$N : \{1,2,...,n\}$:The set of n cities.

$X^i(g,h) = 1$ if there is a direct connection between g and h.
 $= 0$ otherwise.

$X^i(g,h)=X^i(h,g)$

Path: The sequence of cities $(n_1, n_2, ..., n_{z+1})$ is a path if

- (i) $X(n_t, n_{t+1}) = 1$ for $t=1,2,...,z$. and,

(ii) $(n_t, n_{t+1}) \neq (n_s, n_{s+1})$ for $t \neq s$

Two distinct cities (g, h) are said to be connected if there exists a path $(n_1, n_2, \dots, n_z)$ such that $n_1 = g$ and $n_{z+1} = h$.

The size of the network X^i is

$$m^i = \frac{1}{2} \left(\sum X^i(g, h) \right) \quad (96)$$

A network X^i is a hub-and-spoke network if there exists a city h

$$\sum X^i(g, h) = m^i \quad (97)$$

$\pi(z^i, z^j)$ = profit earned by carrier i between cities i and j when there is carrier j in the market. Profit is a non-increasing function of its path.

$$\pi(z, y) = \pi(z + 1, y) \quad (98)$$

Connection function maps a city pair to a path length

$$\Gamma(z) = \{(g, h) \mid \tau(g, h) = z\} \quad (99)$$

The authors assume that:

- 1) Two cities are connected by the shortest path.
- 2) Point-to-point networks are not profitable even if the carrier has a monopoly.
- 3) A carrier cannot earn more than monopoly carriers and,
- 4) No interlining.
- 5) Symmetric network structure.
- 6) Infinite capacity.

Authors assume that profits are purely a function of path length.

$$\pi(z, y) \leq 0, z \geq y$$

Authors then present variety of theorems

Theorem 1)

(i) No pair (X^A, X^B) such that $X^i \in H_m^i$.

$0 < m^i < n-1, i = A, B$ is an equilibrium

(ii) If $X^A \in H_{n-1}$ and $X^B \in \Phi$ then (X^A, X^B) is in equilibrium

Theorem 2) Suppose (Ass1-Ass3) then $\pi(z, y) = 0$ then $z, y < \infty$ and $\pi(z, \infty) = \pi$, then (X^A, X^B) is in equilibrium only if $X^i = \Phi$ and X^j is a tree. This is equivalent to the statement that if both carriers offer paths in the market, then $\pi^A = \pi^B = 0$.

Theorem 3) The best response to a hub-and-spoke network is another hub-and-spoke network

Adler (2001) and Adler (2005) develop a model to evaluate airline profit based on micro-economic theory of behavior under deregulation. Through a two-stage Nash best-response game, equilibria in the air transportation industry is sought to evaluate the most profitable HS network for an airline to survive in a deregulated environment. In the first stage of the game, an integer linear program aids in generating potential networks. In the second stage, a nonlinear mathematical program maximizes profits for each airline, based on the networks chosen by all participants. The variables of the mathematical program include frequency, plane size and airfares. The author then applies this model to a three-airline western European market.

The author assumes that:

- 1) Hub-and-spoke networks reduce total cost to the airline.
- 2) Perceived utility of an airline is proportional to the frequency offered by the airline, in that market.
- 3) Total O-D demand is independent of the carrier's decisions.
- 4) Fares are symmetric and,
- 5) No interlining is allowed.

The model is solved as a two-stage game. In the first stage, the location of hubs is determined using an integer programming formulation. In the second stage the profits of each airline is maximized subject to demand and capacity constraints

The first stage model for hub-location is

Min

$$Z = (\varepsilon_1 + \varepsilon_2)\phi + \sum_j [a_{i1j}x_{i1j} + a_{i2j}x_{i2j}] \quad (100)$$

subject to:

$$x_{i1j} + x_{i2j} = 1, \forall j \neq i1, i2 \quad (101)$$

$$\sum_{j=1}^n x_{i1j} - \sum_{j=1}^n x_{i2j} = \varepsilon_1 - \varepsilon_2, \quad (102)$$

$$x_{i1j}, x_{i2j} \text{ binary} \quad (103)$$

The second stage is a non-linear optimization routine based on game theory

The variables used are

- 1) O-D airfares. (p_{ijsa}).
- 2) Traveler type and airline.
- 3) Aircraft size per leg per airline. ($PS_{h(k,a)}$).
- 4) Frequency per directed leg per airline (f_{ka}).

The model formulation is:

Airline profit = Revenue-Cost

Revenue = $f(\text{Fare, Demand, Market share})$

Cost is expressed as a Cobb-Douglas function of the arc properties

$$Cost = \left[\sum_{k \in Arc} (f_k)^\alpha \right]^\beta \quad (104)$$

Market share = $f(\text{Utility function})$

The utility function is assumed as a multinomial logit function.

$$V_{ijsa} = \lambda_s w_{ijsa} \{\min f_{ka}\}^\gamma - p_{ijsa} \quad (105)$$

Where:

S denotes the type of traveler (business, leisure).

γ is the elasticity of the utility with respect to frequency.

δ_s : willingness to pay for direct flight.

p_{ijsa} : fare from i to j, for passenger type s, for airline a.

$w_{ijsa} = 1 + \delta_s$, for direct flight

=1 for non - direct flight

$$MS_{ijsa} = \frac{\exp(V_{ijsa})}{1 + \exp(V_{ijsa})} \quad (106)$$

Max

$$\sum_{(i,j)} \sum_s \{d_{ijs} MS_{ijsa} \left\langle p_{ijsa} - (PC_i^{nt} + \sum_{k \in R} PC_{c(k,a)}^t) \right\rangle\} - \eta \left(\sum_{k \in R} f_{ka}^\alpha \right)^\beta - \sum_{k \in R} (a_1 + a_2 PS_{h(k,a)}^{SEAT}) LC_{c(k,a)} f_{ka} \quad (107)$$

subject to:

$$L_{ka} \leq PS_{h(k,a)}^{SEAT} \leq U_{ka}, \forall k \in Arc(a), a \in A \quad (108)$$

$$\sum_a \sum_{k \in R} f_{ka} \leq C_i, \forall k \in Arc(a), a \in A \quad (109)$$

$$f_{ka} \geq 0, \forall k \in Arc(a), a \in A \quad (110)$$

$$p_{ijsa} = p_{jisa}, \forall i, j \in N, i \neq j, a \in A \quad (111)$$

Aldregghi, Cento, Nijkamp and Rietveld (2005) identify conditions under which asymmetric equilibria may exist when carriers compete in designing their network configurations in a game-theoretical framework. Two carriers are assumed here, which are allowed to play three different strategies: point-to-point (PP), hub-and-spoke (HS) or multi-hub. The authors execute their model on a four-city, two-airline network.

Game-theoretic models arrive at the equilibrium solution in a single step, unlike supply-demand equilibrium based models. However game-theoretic models suffer from the drawback that they are usually global nonconvex optimization problems, which are very difficult to solve for optimality. Dobson and Lederer (1993) recommend the use of a supercomputer to solve real sized networks consisting for several hundred nodes.

2.1.5) Capacity Constrained Choice Formulations

The logit model approach described in Section 2.1.2 does not take into account capacity constraints on routes; it assumes unlimited capacity on each link. It is possible that a passenger chose a route on account of limited capacity on his preferred route. Ignoring the capacity restriction can lead to a bias in the estimation of route attractiveness.

Soumis and Nagurney (1992) use a stochastic network equilibrium model to estimate the spill on each route, and to model a travelers decision once he has been rejected on a particular route.

Authors develop a methodology similar to dynamic traffic equilibrium to solve the problem on an airline network

The variables used are:

d_m^0 : No of passengers in market m .

Q_m : Set of paths in market m .

p : Path.

X_p : No of passengers selecting path p as their first or subsequent choice.

A_p : Attractiveness of path p (Assuming unlimited capacity).

L_p : Set of arcs in path p .

P_a : Set of paths containing arc a .

f_a : The demand for arc a .

$S_a(f_a)$: The spill on arc a .

$S_p(X)$: The spill on path p .

The model formulation is:

The demand for an arc a is,

$$f_a = \sum_{p \in P_a} X_p \quad (112)$$

Assuming that the spill f_a is random, the expectation of the spill is given by:

$$S_a(f_a) = \sigma_a \frac{e^{-1/2\rho^2}}{\sqrt{2\pi}} + f_a \int_{\rho}^{\infty} \frac{e^{-1/2x^2}}{\sqrt{2\pi}} dx \quad (113)$$

where: $\rho_a = (c_a - f_a)/\sigma_a$.

The spill for path p is therefore,

$$S_p(X) = \sum_{a \in L_p} \frac{x_p S_a(f_a)}{f_a} \quad (114)$$

The realized demand on the market, is the difference of the unconstrained demand and the demand that spilled but not recaptured, which is given by

$$d_m = d_m^0 - \sum_{p \in Q_m} \lambda_p S_p(X) \quad (115)$$

The total flow on path p is given by

$$h_p = X_p - S_p(X)$$

The total number of people that attempt to travel in market m , is

$$u_m = \sum_{p \in Q_m} X_p, \forall m \quad (116)$$

where:

$$X_p = A_p u_m \quad (117)$$

The flow conservation conditions for market m are,

$$\sum_{p \in Q_m} X_p - \sum_{p \in Q_m} (1 - \lambda_p) S_p(X) - d_m^0 = 0 \quad (118)$$

The above set of equations is a non-linear system of simultaneous equations, which are solved by the Gauss-Siedel method. The authors execute their model on the Canadian Air Network. The methodology is explained as a flowchart in Figure 2.3

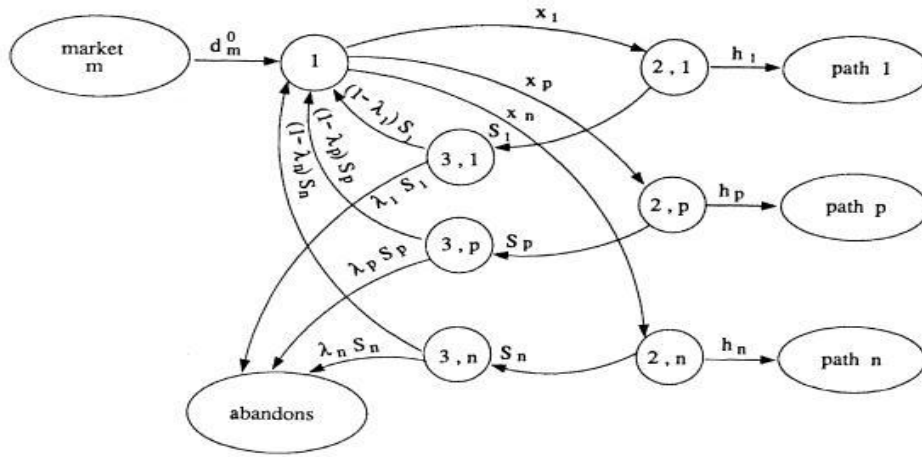


Figure 2.3 Flowchart of the Passenger Decision Making Process in Market m .

2.1.6) Fuzzy Logic Based Formulations

Teodorovic and Kalic (1994) point out that the travelers perceived relative importance of various service attributes cannot be well-defined. Since the route choice model is characterized by uncertainty and subjectivity, the authors develop a model which is based on fuzzy set theory, which utilizes the principles of fuzzy logic. The approximate reasoning model developed is tested using real data.

Authors assume that the route choice model in air transportation is similar to the route choice model in ground transportation. The factors considered in the study were

- 1) Travel Time
- 2) Fares
- 3) Flight frequency
- 4) No of stops

The authors point out that there is a certain amount of fuzziness in these quantities, as perceived by travelers.

The assumptions in the model are:

- 1) If travel costs across routes are similar, then Travel Time and Frequency are the dominant factors in route choice
- 2) Other variables (type of aircraft, seats) are not considered
- 3) Only two alternatives routes exist for every market.

The variables used in the model are:

VBN = Very big negative difference in Travel Time.

BN = Big negative difference in Travel Time.

MN = Medium negative difference in Travel Time.

SN = Small negative difference in Travel Time.

N = Negligible difference in Travel Time.

SP = Small positive difference in Travel Time.

MP = Medium positive difference in Travel Time.

BP = Big positive difference in Travel Time.

VBP = Very big positive difference in Travel Time.

SF = Smaller flight frequency than on alternate route.

ASF = Approximately the same flight frequency than on alternate route.

BF = Bigger flight frequency than on alternate route.

VVS = Very very small percentage of passengers using route A.

VS = Very small percentage of passengers using route A.

S = Small percentage of passengers using route A.

MS = Medium-small percentage of passengers using route A.

M = Medium percentage of passengers using route A.

MB = Medium-big percentage of passengers using route A.

B = Big percentage of passengers using route A.

VB = Very big percentage of passengers using route A.

VVB = Very very big percentage of passengers using route A.

Membership rules are then developed which assign the data to the appropriate fuzzy set.

The rules used in this model are:

Rule 1

If $\delta T = \text{VBN}$ and $\delta F = \text{ANY}$ Then $P = \text{VVB}$

Rule 2

If $\delta T = \text{BN}$ and $\delta F = \text{ANY}$ Then $P = \text{VB}$

Rule 3

If $\delta T = \text{MN}$ and $\delta F = \text{ANY}$ Then $P = \text{B}$

Rule 4

If $\delta T = \text{SN}$ and $\delta F = \text{SF}$ Then $P = \text{M}$

Rule 5

If $\delta T = \text{SN}$ and $\delta F = \text{ASF}$ Then $P = \text{MB}$

Rule 6

If $\delta T = \text{SN}$ and $\delta F = \text{BF}$ Then $P = \text{B}$

Rule 7

If $\delta T = \text{N}$ and $\delta F = \text{SF}$ Then $P = \text{MS}$

Rule 8
 If $\delta T = N$ and $\delta F = ASF$ Then $P = M$
 Rule 9
 If $\delta T = N$ and $\delta F = BF$ Then $P = MB$
 Rule 10
 If $\delta T = SP$ and $\delta F = SF$ Then $P = M$
 Rule 11
 If $\delta T = SP$ and $\delta F = ASF$ Then $P = MS$
 Rule 12
 If $\delta T = SP$ and $\delta F = BF$ Then $P = M$
 Rule 13
 If $\delta T = MP$ and $\delta F = ANY$ Then $P = S$
 Rule 14
 If $\delta T = BP$ and $\delta F = ANY$ Then $P = VS$
 Rule 15
 If $\delta T = VBP$ and $\delta F = ANY$ Then $P = VVS$

The authors test their algorithm on the entire Air Network for the country of Yugoslavia.

The relative magnitude of the fuzzy set variables is calibrated using real data. The advantage of this method, is its computational tractability for large networks. However the model suffers from the drawback that the fuzzy set variables may not be constant across various markets, or between different time-periods. The model also does not account for capacity restrictions.

2.1.7) Heuristic Based Models

Phillips and Boyd (1991) develop a heuristic to spill the passengers along itineraries once their preferred itinerary has reached capacity. The authors present a simple approach for calculating on-board itinerary flows and the corresponding loads, revenues and marginal seat values given a set of underlying demands and an assignment of seating capacities to legs. An application of this approach to a large-scale network is presented and the implication of results for decision making discussed.

The variables used are:

I_{ijk} : k^{th} itinerary from i to j .

$\delta_{ijkm} = 1$ if leg m is a part of I_{ijk}
 $= 0$ otherwise

$f_{ijk}(t)$: fraction of passengers who wish to travel on itinerary I_{ijk}

Therefore the total number of passengers who have reservations on $I_{ijk} = d_{ijk} * f_{ijk}(t)$

$f_{ijk}(t)$: t , assuming constant booking rate over time

$f_{ijk}(0) = 0$, $f_{ijk}(1) = 1$, for all itineraries

The data given to the model is:

Unconstrained itinerary demand, d_{ijk}

C_m : seating capacity of leg m .

Constrained demand is

$L_m : \sum \sum \sum \delta_{ijkm} * L_{ijk}$

K_m : capacity of leg m , not taken up by blocked itineraries (Blocked itineraries = at least one leg in that itinerary has become full)

$U_{ijk} = 0$ if I_{ijk} is blocked

$= 1$ otherwise

The methodology adopted by the authors is:

Step 0 Initialize

$U_{ijk} = 1$, for all itineraries

$K_m = C_m$ for all legs

$l_{ijk} = 0$ for all itineraries

$t = 0$

Step 1 For each leg calculate

$r_m = \sum \sum \sum \delta_{ijkm} * U_{ijk} * d_{ijk}$

$t_m = K_m / R_m$

Step 2

For the system determine $m^* = \arg \min(t_m)$

m is the next leg that becomes full.

if $t_m \geq 1$ Go to step 5

$t_m < 1$ Go to step 3

Step 3

For all i, j, k such that $\delta_{ijkm} = 1$ and $U_{ijk} = 1$

$L_{ijk} = T_m * d_{ijk}$

$U_{ijk} = 0$

If $U_{ijk} = 0$ for all i, j, k go to step 5

Step 4

$k_m = C_m - \sum \sum \sum \delta_{ijkm} * L_{ijk}$ for all m

Go to step 1

Step 5

Set all $L_{ijk} = L_{ijk} + U_{ijk} * d_{ijk}$

The advantage of this model is the ease of computation of the itinerary flows. However the model suffers from the weakness, that it requires information about booking curves and itinerary demand which is too microscopic for the Air Network Flow Problem.

2.1.8) Probability Based Formulations

Jansson and Ridderstolpe (1992) develop a model to determine passenger route choice in public transit systems. The authors argue that passengers in an urban transport system may often be in a position to choose among a set of parallel routes in order to reach their destinations. Their decisions depend in general on many variables e.g., the in-vehicle times, walk time, transfer time, number of transfers, stops and the headways. Traditional methods used in urban transport planning do not solve the route choice problem satisfactorily when passengers may choose between modes with differences in speed or fare structures e.g., tram and bus

The authors assume that the passengers are of two types

Type A: The passengers know the riding times, the headways and the departure times associated with all the routes

Type B: The passengers know the riding times, the headways but do not know the departure times associated with all the routes

The operational assumption concerning the co-ordination of routes is classified as type C or D.

Type C: The interval between departure times on any given route is constant. Departures of different routes are independent.

Type D: The interval between departure times on any given route is constant. Departures of different routes are perfectly coordinated.

The variable used are:

L : The set of parallel routes.

p : Ridership shares.

r : Riding times.

w : Average waiting times.

t : Headways.

The probability p of taking route u and the average waiting time w is

$$p_u = \int_0^{t_u} \frac{1}{t_u} \prod_{i=1}^k (1 - \min\{1, \max\{\frac{x + r_u - r_i}{t_i}, 0\}\}) \quad (119)$$

$$w = \sum_{u=1}^k \int_0^{t_u} \frac{1}{t_u} \prod_{i=1}^k (1 - \min\{1, \max\{\frac{x + r_u - r_i}{t_i}, 0\}\}) \quad (120)$$

The expressions for both these parameters are not computationally tractable for networks greater than two parallel routes. The authors use this property of the formulation for their solution methodology. The solution is based on a successive chain of choices between a route and an "equivalent" route, representing by a cluster of routes. Accordingly one route at a time is compared with a set of already examined and accepted routes. If the route has feasible attributes, it is included in the acceptable set of routes, otherwise it is rejected. The waiting times and path distribution are calculated in an iterative manner, in which at each step a route is compared with a cluster of incumbent routes which form the "equivalent route".

Rules employed in the heuristic

Rule 1. The routes are processed in ascending order of riding time (the fastest first)

Rule 2. The currently examined route u is accepted only if its riding time is less than riding time plus headway for an "equivalent" route m composed of all routes with shorter riding times. i.e. $r_u < r_m + t_m$.

Rule 3. The shares of the routes already accepted will be reduced by a factor of $1 - p_u$, where p_u is the share of the currently added route.

The advantage of the method is its computational tractability. The disadvantage is that factors like fares, congestion are not taken into account. Capacity constraints along the routes are also not taken into account.

Table 2.1 Summary of the Methods used for the Air Traffic Network Flow Problem.

Methodology	Benefits	Drawbacks	Characteristics	Size of the Problem
Entropy	Does not depend on Network attributes	No causality is modeled	Non-linear (logarithmic) Programming	5000 node, 10000 link problem
Logit based formulations	Causality is modeled	No supply side variables considered	Maximum likelihood	125 hub airports, 7405 markets
Supply-demand equilibrium formulations	Both supply and demand side variables are modeled	Iteration is required for convergence	Maximum likelihood	31 top hub airports, Taiwan domestic air network.
Supply-constrained logit model	Both supply and demand side variables are modeled	Iteration is required for convergence	Non-linear simultaneous equations	The Canadian air network
Game-theoretic models	Models competition, demand-supply interaction	Difficult to solve for large networks	Non-linear optimization	Five city, three airline problem.
Fuzzy logic based models	Computationally simple	Rules are not clearly defined and fuzzy sets can	Fuzzy set theory, assume uncertainty in decision process	Yugoslav air network
Heuristic based formulations	Computationally simple	Requires microscopic data	Assumes rules for flowing passengers	The network for United Airlines
Probability based formulations	Computationally simple	No capacity constraints	Route probability function	Six routes

2.2) Literature Review on the Schedule Generator Problem

The objective of the Schedule Generator Problem is to determine the frequency of flights, the arrival and departure time of each flight, its origin and destination and the aircraft type used to operate the flight. Flight schedules are of two types

- (1) Periodic schedules, which repeat every day without any changes
 - (2) Non-periodic schedules, which undergo minor changes depending on the demand.
- Airlines typically do not operate periodic schedules.

This research does not take into account demand changes across different days of the week, or different months of the year. A single daily schedule is constructed which is then assumed to repeat every day for the entire year. The Schedule Generator problem can be formally defined as:

Given a set of Origin-Destination pairs, the arcs covering each O-D pair, the demand for each arc, the fare paid by the passengers in the market, the fleet type and the operating cost for operating an aircraft in the fleet

Find the schedule set that minimizes the airline operating cost, and satisfies demand and capacity constraints

Various techniques and models have been proposed in literature to derive the optimal schedule. They can be sub-divided into four categories

- (1) Mathematical programming formulations, independent of previous schedules.
- (2) Mathematical programming formulations, based on previous schedules.
- (3) Supply-demand based formulations.
- (4) Heuristic based-formulations.

2.2.1) Mathematical Programming Formulations, Independent of Previous Schedules

Models that build a new schedule that is independent of previous schedules are used principally for air taxi operations, which do not follow regular operations unlike commercial airlines. However the same approach has also been used to derive optimal schedules for regular scheduled carriers.

One of the first models to optimize flight operations was proposed by Teodorovic (1983) . He developed a model to determine the number of flights per day between two cities, assuming direct flights and no competition. The author measures the economic effects of using a particular flight frequency using the load factor value. The level of service is measured by the absolute value of the average time difference between the actual and desired time of departure. It is shown that this time difference can be approximately expressed only as a function of flight frequency, without regard to the times of departure

in the course of the day and that the error which occurs is not considerable. The interval in which the flight frequency should be located is also determined, as well as the interval of the average time difference between the actual and desired time of departure in order to achieve a profitable business.

The variables used in the model are:

T : Time interval considered.

$x_1, x_2, \dots, x_n$: Time of departure of flights.

$h(t)$: Time dependent demand function.

The departure x_i will be chosen by all the passengers who want to travel in the interval

$$\left[\frac{x_{i-1} + x_i}{2}, \frac{x_i + x_{i+1}}{2} \right]. \quad (121)$$

$$H(t) = \int_0^t h(t) dt$$

W : Total time difference between actual and desired departure time for all passengers, in the interval $[0, T]$ is:

$$\begin{aligned} W = & 2 \int_0^{x_1} H(t) dt + 2 \int_{x_1}^{x_2} H(t) dt + \dots + 2 \int_{x_{N-1}}^{x_N} H(t) dt - 2 \int_0^{(x_1+x_2)/2} H(t) dt - 2 \int_0^{(x_2+x_3)/2} H(t) dt - \dots - 2 \int_0^{(x_{N-1}+x_N)/2} H(t) dt \\ & - x_N H(T) - \int_0^T H(t) dt + TH(T) \end{aligned} \quad (122)$$

Where $H(T) = Q$, the average number of passengers requesting service in $[0, T]$.

The authors then calculate the departure times, by setting $dW/dx = 0$.

The departure times are:

$$H(x_1) = \frac{H\left(\frac{x_1 + x_2}{2}\right)}{2} \quad (123)$$

$$H(x_2) = \frac{H\left(\frac{x_1 + x_2}{2}\right) + H\left(\frac{x_2 + x_3}{2}\right)}{2} \quad (124)$$

$$H(x_3) = \frac{H\left(\frac{x_2 + x_3}{2}\right) + H\left(\frac{x_3 + x_4}{2}\right)}{2} \quad (125)$$

However, airline fleet size is limited, which constrains the profitability of the schedule. Researchers started to look at models, which generated a realistic schedule that took the fleet size of the aircraft into account.

Teodorovic (1987) describes a model to solve the aircraft assignment and schedule generation problem simultaneously, which considers fleet size. The author makes the following assumptions to simplify the problem:

- 1) Only non-stop flights are considered.
- 2) The level of service is related to schedule delay alone.

Then flights are assigned to routes based on the following criterion

- 1) Capital investment.
- 2) Average operating costs per passenger.
- 3) Percentage of routes in the network having daily flights.
- 4) Average schedule delay per passenger.
- 5) Number of aircraft in the fleet.

The weights for each alternative are determined using the following method:

- 1) Consider each permutation of the alternatives. If there are m alternatives there are $m!$ permutations.

- 2) For each permutation,

Calculate

$$R_i = \sum_{Cst} W_j - \sum_{Dst} W_j$$

Where

$$C_{st} = \{j | x_{sj} \geq x_{st}\}$$

$$D_{st} = \{j | x_{sj} \leq x_{st}\}$$

The permutation with the highest value of R_i is taken and the order of importance determined.

- 3) Once the order of importance of alternatives is known, the weights are determined from the relation:

$$\sum_j W_j = 1$$

Such that

$$W_s \geq W_t$$

The author then proposes a mathematical program to minimize schedule delay while satisfying fleet constraints:

$$SD = \frac{Q_i + 2N_i}{4Q_i N_i} T \quad (126)$$

Q_i = average number of passengers that wish to travel on route i in $[0, T]$

N_i = flight frequency on route i in $[0, T]$

$$W = \sum_{r=1}^p r_i Q_i = \frac{pT}{2} + \frac{T}{4} \sum_{r=1}^p \frac{Q_i}{N_i} \quad (127)$$

The average delay per passenger

$$w = \frac{pT}{2 \sum_{r=1}^p Q_i} + \frac{T}{4 \sum_{r=1}^p Q_i} \sum_{r=1}^p \frac{Q_i}{N_i} \quad (128)$$

The frequency is determined from,
Min

$$w = \frac{pT}{2 \sum_{r=1}^p Q_i} + \frac{T}{4 \sum_{r=1}^p Q_i} \sum_{r=1}^p \frac{Q_i}{N_i} \quad (129)$$

subject to:

$$\sum_{r=1}^p N_i n_{Bi} \leq S \quad (130)$$

Following deregulation, airlines networks grew rapidly and one route was, in many cases served by more than one airline. Researchers explored models to derive the optimal schedule when more than one carrier were operating along a route.

Teodorovic (1989) developed to determine flight frequencies on a route network when there is competition. In determining flight frequencies the focus is placed on maximizing profit, maximizing the number of passengers flown and minimizing the total passenger schedule delay. The authors assume that,

- 1) All the carriers operate the same aircraft type in a route.
- 2) The carriers charge the same ticket price on a route.

The market share of airline p on route i can be written as:

$$MS_{ip} = \frac{N_{ip}^\alpha}{\sum_j N_{ij}^\alpha} \quad (131)$$

Where:

MS_{ip} = market share on route i for airline p .

α = constant ($1 < \alpha < 2$).

The number of passengers carried by airline p on route i can be written as,

$$V_{ip} = \mu_i MS_{ip}.$$

μ_i = Total demand on route i .

The profit is,

$$P_{ip} = c_i \mu_i MS_{ip} - C_i N_{ip} \quad (132)$$

The total profit of airline p is obtained by summing the profit across all routes

$$P_p = \sum_i c_i \mu_i \left(\frac{N_{ip}^\alpha}{\sum_j N_{ij}^\alpha} \right) - C_i N_{ip} \quad (133)$$

The total schedule delay experienced by passengers in the network of carrier p is,

$$SD_p = \sum_i V_{ip} SD_{ip} = \sum_i \frac{V_{ip} T}{4N_{ip}} \quad (134)$$

The total passengers carried by airline p can be written as:

$$T_p = \sum_p V_{ip} \quad (135)$$

This can be rewritten as:

$$T_p = \sum_i \mu_i \left(\frac{N_{ip}^\alpha}{\sum_j N_{ij}^\alpha} \right) \quad (136)$$

The constraints are:

- 1) The frequency of an airline on a route cannot exceed the maximum allowable frequency ($N_{ip} \leq N_{ip}^*$).
- 2) The total passenger seats carried by airline cannot exceed its seat capacity ($\sum N_{ip} n_{tBi} \leq S_p$).
- 3) The airline would like to avoid losses on some routes ($C_i N_{ip} \leq c_i \mu_i M S_{ip}$).

The problem of determining the flight frequency can be formulated as multi-objective optimization problem:

Max

$$P_p = \sum_i c_i \mu_i \left(\frac{N_{ip}^\alpha}{\sum_j N_{ij}^\alpha} \right) - C_i N_{ip} \quad (137)$$

Max

$$T_p = \sum_i \mu_i \left(\frac{N_{ip}^\alpha}{\sum_j N_{ij}^\alpha} \right) \quad (138)$$

Min

$$SD_p = \sum_i V_{ip} SD_{ip} = \sum_i \frac{V_{ip} T}{4N_{ip}} \quad (139)$$

subject to:

$$N_{ip} \leq N_{ip}^*, \forall i, p \quad (140)$$

$$C_i N_{ip} \leq c_i \mu_i MS_{ip}, \forall i, p \quad (141)$$

$$N_{ip} nt_{Bi} \leq S_p, \forall i, p \quad (142)$$

The authors use a Monte Carlo approach to arrive at the optimal solution.

As airlines expanded their operations, it became necessary to consider more factors into the schedule planning operations, and construct more robust models that optimized the schedule design while considering multiple factors. It also became imperative to consider network system effects since changing the attributes of one leg, has a cascading effect on the whole network.

Yan (1996) developed a model that maximizes system profit when determining aircraft frequencies and aircraft routes in a network. The model employs network flow techniques to collect or deliver passengers from all origins to all destinations using non-stop and multi-stop flights in multi-fleet operations. The author assumes that each carrier has a maximum limit on the number of operations per airport to satisfy slot criteria

The inputs to the model are:

- Demand (weekly)
- Aircraft types
- Operational Cost/Revenue data
- Frequency constraints

The model formulation is:

Variables

ϕ_{ij} : Quota for segment (i,j) .

p_{ij}^k : Arc cost in (i,j) for aircraft type k .

y_{ij}^k : Aircraft flows in (i,j) for aircraft type k .

C_{ij}^n : Arc cost for the n th OD pair on segment (i,j) .

x_{ij}^k : Passenger flows in (i,j) for aircraft type k .

U_{ij}^n : Upper bound for arc (i,j) in passenger network n .

Minimize

$$\sum_n \sum_{ij} C_{ij}^n X_{ij}^n + \sum_k \sum_{ij} P_{ij}^k Y_{ij}^k \quad (143)$$

subject to:

$$\sum_j X_{ij}^n - \sum_j X_{ji}^n = 0, \forall i \quad (144)$$

$$\sum_j Y_{ij}^n - \sum_j Y_{ji}^n = 0, \forall i \quad (145)$$

$$\sum_n X_{ij}^n \leq \sum_k Y_{ij}^k U^k \eta_{ij}, \forall i, j \quad (146)$$

$$\sum_i \sum_k w^k y_{ij}^k \leq \phi_j, \forall j \quad (147)$$

$$\sum_k w^k y_{ij}^k \leq \phi_{ij}, \forall i, j \quad (148)$$

$$\sum_j \sum_k w^k y_{ij}^k \leq \phi_i, \forall i \quad (149)$$

$$U_{ij}^n \geq X_{ij}^n \geq 0, \forall i, j \quad (150)$$

The author uses a lagrangian based method to solve the above integer program.

Desaulniers, Desrosiers and Dumas (1997) propose a model to solve the daily aircraft routing and scheduling problem, which is defined by the authors as follows

Given a heterogeneous aircraft flight fleet and set of flight legs over a one day horizon, departure time windows, duration and costs according to the aircraft type for each leg, find a fleet schedule that maximizes profits and satisfies certain additional constraints.

The authors use two concurrent approaches:

- 1) A set partitioning approach with assigns a binary variable for each possible schedule of the aircraft.
- 2) A binary variable for each connection performed by a particular aircraft (Time-constrained multi commodity network flow).

The variables used by the authors are:

N : The set of operational flight legs.

n^k : The number of available aircraft of type k .

Ω^k : The set of feasible schedules for aircraft k .

C_p^k : Profit from assigning aircraft k to schedule p .

$a_{ip}^k = 1$ if schedule p by aircraft k covers flight leg i .
 $= 0$ otherwise

S : The set of stations.

S^k : The set of stations served by k .

$O_{sp}^k = 1$ if schedule p starts at station p .
 $= 0$ otherwise.

$d_{sp}^k = 1$ if schedule p ends at station p .
 $= 0$ otherwise

$\theta_p^k = 1$ if schedule p is assigned to an aircraft of fleet type k .
 $= 0$ otherwise

The authors discuss a method to generate the set of feasible schedules:

A network can be represented as $G^k = (V^k, A^k)$, where V^k is the set of nodes and A^k is the set of arcs.

The set V^k consists of

- 1) Source.

- 2) Sink.
- 3) Initial station.
- 4) Final station.
- 5) Flight node.

The set A^k consist of

- 1) Empty (The arc connects the source and the sink nodes)
- 2) Source (The source arc connects the source node to the node representing the initial station)
- 3) Sink (The sink arc connects the destination node to the node representing the final station)
- 4) Schedule start (The schedule start arc connects the initial station node to the node representing the start of the schedule)
- 5) Schedule end (The schedule end arc connects the final station node to the node representing the end of the schedule)
- 6) Turn arc (The turn arc represents a normal connection between two flights)

Let a_i^k = earliest time leg i can start.

b_j^k = latest time leg j can start.

t_{ij}^k = normal connecting time between i and j .

l_i^k = duration of leg i .

Then,

$$a_i^k + l_i^k + t_{ij}^k \leq b_j^k$$

- 7) Short turn arc: The short turn arc represents a short connection between two flight legs

Let a_i^k = earliest time leg i can start.

b_j^k = latest time leg j can start.

s_{ij}^k = short normal connecting time between i and j .

l_i^k = duration of leg i .

$$a_i^k + l_i^k + s_{ij}^k \leq b_j^k$$

The Cost (Profit) of a schedule is: $r_i^k - e_i^k$

X_{ij}^k : Integer flow variable on arc (i,j) for aircraft type k .

T_i^k : Time variable defined on a node i , T_i^k belongs to $[a_i^k, b_i^k]$.

The model formulations are:

Formulation I (The set partitioning approach)

Max

$$Z = \sum_{k \in K} \sum_{p \in \Omega^k} C_p^k \theta_p^k \quad (151)$$

subject to:

$$\sum_{k \in K} \sum_{p \in \Omega^k} a_{ip}^k \theta_p^k = 1, \forall i \quad (152)$$

$$\sum_{p \in \Omega^k} (d_{sp}^k - o_{sp}^k) \theta_p^k = 0, \forall s, k \quad (153)$$

$$\sum_{p \in \Omega^k} \theta_p^k = n^k, \forall k \quad (154)$$

Formulation II (Time-constrained multi commodity network flow approach)

Max

$$\sum_{k \in K} \sum_{(i,j) \in A^k} C_{ij}^k X_{ij}^k \quad (155)$$

subject to:

$$\sum_{k \in K} \sum_{j: (i,j) \in A^k} X_{ij}^k = 1, \forall i \quad (156)$$

$$\sum_{i: (i,s) \in N, S_2^k} X_{is}^k - \sum_{j: (s,j) \in N, S_1^k} X_{sj}^k = 0 \quad (157)$$

$$\sum_{i: (i,s) \in S_1^k} X_{o(k),s}^k - X_{o(k),d(k)}^k = n^k, \forall k \quad (158)$$

$$\sum_{i: (i,s) \in S_2^k} X_{s,d(k)}^k - X_{o(k),d(k)}^k = n^k, \forall k \quad (159)$$

$$\sum_{i: (i,j) \in A^k} X_{ij}^k - \sum_{j: (j,i) \in A^k} X_{ij}^k = 0, \forall k \quad (160)$$

Braanlund (1998) proposes a unique way to profitably schedule trains, while considering track capacity constraints. This problem is analogous to deriving the profit maximizing schedule of an airline, considering airport and leg capacity constraints.

The problem formulation is:

Let $x_{it}^r = 1$ if train r occupies block i at time t .
 $= 0$ otherwise.

Let x^r be the corresponding vectors of x_{it}^r 's. Let T^r be the set of x^r 's denoting the technically feasible, logical feasible schedule for train r .

The profit maximization problem can be written as

Max

$$\sum_r V^r(X^r)$$

subject to:

$$X_{it}^r = 1 \forall i, t, X^r \in T^r$$

Erdmann and Noltemeier (1999) use a binary integer programming formulation to schedule an air taxi service. The authors propose a combined Branch-and-Cut approach to solve the airline schedule generation problem. To tighten the linear relaxation bound, they add cutting planes which adjust the number of aircraft and the spill of passengers to the demand on each itinerary. For real-world problems, using data from a large European charter airline they obtain solutions within a very few percent of optimality with running times in the order of minutes.

The variables used are:

A : The set of all airports.

K : The set of all aircrafts indexed by k .

Ω_d^k : The set of feasible rotations for aircraft k on day d , indexed by p .

M : The set of all O-D pairs.

I_d^m : The set of all itineraries for OD pair m , on day d .

Decision variables

$\theta_p^k = 1$ if rotation p is flown by aircraft k .
 $= 0$ otherwise

Y_{ad}^k : slack/surplus variable to model the position of aircraft k on airport a .

x_i : no of passengers taken on itinerary i .

s_d^m : The passenger spill on OD pair m , on day d .

Data

N : The length of the planning period in days.

C_p^k : The cost for using aircraft k on rotation p .

C_i : The cost per passenger flown for itinerary i .

Cap^k : The seat capacity of aircraft k .

$n_d^k = 1$ if aircraft k is available on day d .
 $= 0$ otherwise

r_{ad}^k, s_{ad}^k : lower and upper bounds for Y_{ad}^k .

dem_d^k : Demand for O-D pair m on day d .

$\delta_{fp}, \delta_{fi} = 1$ if rotation p or itinerary i have aircraft f .
 $= 0$ otherwise

$\delta_{ip} = 1$ if the itinerary i and aircraft rotation p are sharing a flight.

$O(i), D(i)$: The origin and destination for passenger i .

$O(p), D(p)$: The origin and destination for rotation p .

$O(f), D(f)$: The origin and destination for flight f .

The authors assume that

- 1) Each aircraft has associated home base where it is parked overnight.
- 2) Demands are symmetric.
- 3) Capacities are symmetric.
- 4) Each itinerary has a maximum of three legs.

The model formulation is:

Min

$$Z = \sum_{1 \leq d \leq N} \left(\sum_{k \in K} \sum_{p \in \Omega} C_p^k \theta_p^k + \sum_{m \in M} \sum_{i \in I} C_i x_i \right) \quad (161)$$

subject to:

$$\sum_{m \in M} \sum_{i \in I_d^m} \delta_{fi} x_i \leq \sum_{k \in K} cap^k \sum_{p \in \Omega_d^k} \chi_{fp}^k \theta_{fp}^k, \forall f \quad (162)$$

$$\sum_{p \in \Omega_d^k} \theta_p^k \leq n_d^k, \forall d, k \quad (163)$$

$$\sum_{p \in \Omega_{d+1}^k} \theta_p^k - \sum_{p \in \Omega_d^k} \theta_p^k - Y_{ad}^k = 0, \forall a, d, k \quad (164)$$

$$\sum_{p \in \Omega_1^k} \theta_p^k - \sum_{p \in \Omega_N^k} \theta_p^k - Y_{aN}^k = 0, \forall a, d, k \quad (165)$$

$$r_{ad}^k \leq Y_{ad}^k \leq S_{ad}^k, \forall a, d, k \quad (166)$$

$$\sum_{i \in I_d^m} X_i + s_d^m \leq dem_d^m, \forall m, d \quad (167)$$

$$\theta_p^k \text{ binary}, x_i \geq 0, s_d^m \geq 0 \quad (168)$$

The models described above derive the optimal schedule without taking into account any aircraft maintenance constraints. As airline networks grew, and fleet utilization grew, and FAA maintenance guidelines became more stringent, airlines felt the need to optimize schedules and maintenance operations simultaneously. Feo and Bard (1990) propose a model for the simultaneous design of schedules and maintenance bases.

The variables used are:

n_d : Planning horizon (days).

n_c : The number of cities.

n_p : The number of aircraft in the fleet.

j, k : The index for cities ($j, k = 1, 2, \dots, n_c$).

$x_{ij(d)k(d+1)} = 1$, if plane i is in city j on the evening of day d and in city k on the evening of day $d+1$
 $= 0$, otherwise

$\delta_j = 1$, if j is a base.

$= 0$, otherwise.

y_{id} : The number of days before plane i gets maintenance.

g_j : The cost of maintenance in city j .

E : The set of edges of the graph, as defined in the schedule.

The authors point out that the problem is a large scale MIP (10^6 variables) and propose a few simplifications

- 1) The problem is separated into different sub problems for various plane types
- 2) Base capacity constraint is eliminated

Min

$$Z = \sum_{j=1}^{n_c} [f_j \delta_j + \sum_{i=1}^{n_p} \sum_{d=1}^{n_d} g_j w_{ij(d)}] \quad (169)$$

subject to:

$$\sum_{j(d-1)} X_{ij(d-1)k(d)} - \sum_{j(d+1)} X_{ik(d)j(d+1)} = 0, \forall i, k(d) \quad (170)$$

$$\sum_{i=1}^{n_p} X_{ij(d)k(d+1)} = 1, \forall d, (j(d), k(d+1)) \in E \quad (171)$$

$$\sum_{i=1}^{n_p} w_{ij(d)} \leq p_j \delta_j, \forall d, j(d) \quad (172)$$

$$w_{ij(d)} - \sum_{k(d-1)} X_{ik(d-1)j(d)} \leq 0, \forall i, j(d) \quad (173)$$

$$Y_{i,d+1} - Y_{id} - 4(\sum_{j(d)} w_{ij(d)}) \leq -1, \forall i, d \quad (174)$$

$$X_{ij(d)k(d+1)}, w_{ij(d)}, \delta_j \text{ binary} \quad (175)$$

The schedule design algorithms have also been successfully applied to optimize the operations of cargo carriers. Yan, Lai and Chia (2005) present a model to design the schedule of a short-term freight carrier service. The model is formulated as an integer multiple commodity network flow problem and solved using mathematical programming.

The authors use a space-time network to model a schedule for cargo delivery. Separate time-space networks are built for each fleet type. The following arcs are considered:

Flight arcs - Flight arcs are either pickup flight arcs or delivery flight arcs. All reasonable flight arcs are considered. The upper bound for each flight arc is set to 1 (an aircraft is assigned to the arc) or 0 (the arc is not assigned any aircraft).

Ground arcs - Represents the holding or the overnight stay of an aircraft at an airport. The upper bound for arc capacity is the apron capacity of the airport and the lower bound is zero.

Cycle arcs - Cycle arcs connect the beginning of one cycle to the end of the previous cycle. The capacity constraints for cycle arc is same as that of the ground arc.

Alternate flight arcs - An alternate flight arc represents a slight perturbation of an existing flight arc. The capacity constraints are same as that of a regular flight arc.

The variables used are:

- n – The number of fleets used.
- M – The set of all fleets.
- N^n – The set of all arcs in the n^{th} fleet network.
- A^n – The set of ground arcs in the n^{th} fleet network.
- A_f^n – The set of flight arcs in the n^{th} fleet network.
- A_r^n – The set of cycle arcs in the n^{th} fleet network.
- F^n – The number of available aircraft in the n^{th} fleet network.
- B_a – The set of alternative flight arcs for flight a .
- b_i – The demand/supply of node i .
- U_{ij}^n – The upper bound for arc flow in the n^{th} fleet network for arc (i, j) .
- C_{ij}^n – The cost for arc flow in the n^{th} fleet network for arc (i, j) .
- X_{ij}^n – The actual arc flow in the n^{th} fleet network for arc (i, j) .

The model formulation is

Min

$$Z = \sum_{n \in M} \sum_{ij \in A} C_{ij}^n X_{ij}^n \quad (176)$$

subject to:

$$\sum_{j \in N^n} X_{ij}^n - \sum_{k \in N^n} X_{ki}^n = b_i, \forall i \in N^n, n \in M \quad (177)$$

$$\sum_{ij \in A^n} X_{ij}^n \leq F^n, \forall n \in M \quad (178)$$

$$\sum_{n \in M} \sum_{ij \in B^a} X_{ij}^n = 1 \quad (179)$$

$$\sum_{ij \in R_1^c} X_{ij}^n - \sum_{kl \in R_2^c} X_{kl}^n = 0, \forall n \in M, c \in S \quad (180)$$

$$\sum_{n \in M} \sum_{ij \in A_f^n} X_{ij}^n + \sum_{n \in M} \sum_{ik \in A_f^n} X_{ik}^n \leq Q_i, \forall i \in H \quad (181)$$

$$0 \leq X_{ij}^n \leq U_{ij}^n, ij \in A^n, n \in M \quad (182)$$

$$X_{ij}^n \text{ integer} \quad (183)$$

As flights volumes increased, the delays at airports particularly at major hubs were exacerbated. Researchers were confronted with the problem of constructing a reliable airline schedule, i.e. a schedule that is resistant to disruptions (weather phenomena, hub closures).

Rosenberger, Johnson and Nemhauser (2004) describe a model to construct a flight schedule that is resistant to cancellations.

The variables used are:

J : The set of fleets.

For each j in J , let

$P(j)$: No of aircrafts in j .

κ_j : Capacity of the aircraft in j .

τ_j : Cost per hour of operating an aircraft in fleet type j .

F : The set of flight legs.

m_f : The passenger demand for leg f .

v_f : Average ticket price.

$S(f)$: The set of strings that include flight leg f .

b_f : The block time for flight leg f .

S : The set of strings.

N : The set of nodes in the timeline network.

$X_{js} = 1$ if fleet type j is assigned to string s .

$= 0$ otherwise

C_{js} : The cost of assigning fleet type j to string s

$$= \sum_f v_f (m_f - \kappa_j) + b_f \tau_j$$

The authors model the aircraft operations using “flight strings”. Flight strings are a small sequence of legs, flown by a fleet type .The authors propose the idea that flight strings that do not originate and end in the same hub, are the ones that would be most affected by a disruption, since they would be unable to return to their base at the end of the string.

The asymmetry of the flight strings, is encapsulated in a variable called hub connectivity which is the proportion of the strings whose origin and destination hubs are different. The proposed model tries to minimize the hub-connectivity of a network, and consequently mitigating the effect of a hub disruption on the network. The set of feasible strings considered by the model are symmetric strings. The model formulation is:

Min

$$Z = \sum_{s \in S} C_{js} f_{js} \quad (184)$$

subject to:

$$\sum_{s \in S} \sum_{j \in J} X_{js} = 1 \quad \forall f \in F \quad (185)$$

$$\sum_{s \in O(j)} X_{js} + \sum_{g \in G} Y_{jgtnt1} \leq P(j), \quad \forall j \in J \quad (186)$$

$$\sum_{ogt_o, t \in S} X_{jogt_o t} + Y_{jgt^{-}t} - \sum_{gdt_o, t} X_{jgtt_f} - Y_{jgtt^{+}} = 0, \quad \forall jgt \in N \quad (187)$$

$$Y_{jgtt^{+}} \geq 0, \quad \forall jgt \in N \quad (188)$$

$$X_{js} \in \{0,1\}, \quad \forall s \in S, j \in J \quad (189)$$

Mathematical programming formulations that construct a profit-maximizing schedule are well-suited to air taxi and other irregular carriers, where the schedules have significant degree of freedom. On the other hand, the models presented above may not mirror the real schedules of the airline industry. Therefore, it may be desirable that current schedules of the air carriers be taken into account, while designing new schedules.

2.2.2) Mathematical Programming Formulations, Dependent on Previous Schedules

Schedule generation algorithms, which utilize previous (base) schedules and design new schedules by incremental changes in the base schedules, are a recent development. These studies were motivated by the need to redesign the schedule in real-time, following a schedule perturbation. Since the redesigned schedule can not be very different from the base schedule, the need arose to design new schedules by tweaking the old one.

Yan (1996) develops a framework to assist carriers in fleet routing and flight scheduling for schedule perturbations in the operations of multifleet and multistop flights. The framework is based on a basic multifleet schedule perturbation model constructed as a timespace network. The authors develop a space-time network based model for schedule perturbations. They start on the assumption that an aircraft is unavailable at time $t=0$. The model starts from $t=0$ and ends when the schedule returns to normal again. The time-space network has four kinds of arcs:

- 1) Flight arcs: Represents a full passenger flight between two stations. The upper bound for the arc capacity is set to 1 and lower bound is set to zero. Arc cost is written as: $C_{cij} + (C_{ij} - C_{cij}) * X_{ij}$
Where C_{ij} =arc flow cost
 C_{cij} =flight cancellation cost.
 $X_{ij} = 1$ if a flight is made between i and j
 $= 0$ otherwise
- 2) Ground arc: A ground arc represents the time when an aircraft is held on the airport. The arc cost is the cost that an airport charges to hold an aircraft. The capacity of the arc is taken as the number of airplanes that can be accommodated on the apron.
- 3) Overnight arc: This arc represents the event of an aircraft being held in an airport overnight. The cost of an overnight arc is the same as that of the ground arc.
- 4) Ferry arc: The ferry arc represents the time it takes to reposition an aircraft. The arc capacity is taken as the arrival capacity of the airport. The arc cost is the combination of fuel and crew cost and lost passenger revenue.

The variables used in the model are:

n : The number of aircraft types.

M : The set of all aircraft types.

s : The node where and when an aircraft is absent from the system.

t : The node where and when the absent aircraft is ready for service.

SF : The set of all flights.

N^n : The set of all nodes in the n^{th} type network.

A^n : The set of all arcs in the n^{th} type network.

A_f^n : The set of all flight arcs in the n^{th} type network.

$C_{ij}^n, X_{ij}^n, U_{ij}^n$: The arc cost, actual flow and upper bound flow for arc (i,j) .

C_{cij}^n : The cancellation cost for a flight (i,j) in the n^{th} flight network.

$C_{dij}^n : C_{ij}^n - C_{cij}^n$.

Δ : The number of absent airplanes.

The model formulation is:

Min

$$Z = \sum_{n \in M} (C_{cij}^n + \sum_{i,j \in A_f^n} C_{dij}^n X_{ij}^n + \sum_{i,j \in A / A_f^n} C_{ij}^n X_{ij}^n) \quad (190)$$

subject to:

$$\sum_{j \in O(s)} X_{sj}^p - \sum_{k \in L(s)} X_{ks}^p = b_s^p - \Delta \quad (191)$$

$$\sum_{j \in O(t)} X_{tj}^p - \sum_{k \in L(s)} X_{ks}^p = b_t^p - \Delta \quad (192)$$

$$\sum_{j \in O(i)} X_{ij}^p - \sum_{k \in L(i)} X_{ki}^p = b_i - \Delta \quad (193)$$

$$\sum_{n \in M} X_{ij}^n \leq 1, \forall (i, j) \in SF \quad (194)$$

$$0 \leq X_{ij}^n \leq 1, \forall (i, j) \in A_f^n, \forall n \in M \quad (195)$$

$$0 \leq X_{ij}^n \leq U_{ij}^n, \forall (i, j) \in A^n / A_f^n, \forall n \in M \quad (196)$$

$$X_{ij}^n \text{ integer}, \forall (i, j) \in A^n, \forall n \in M \quad (197)$$

Models to construct new schedules in real-time upon perturbation in an old schedule must be robust and computationally tractable simultaneously. This is achieved by restricting the degrees of freedom of change that is allowable in the old schedule.

Teodorovic and Stojkovic (1990) discuss a way of quickly designing a new schedule once perturbation has been made to an old schedule.

The variables used in the model are:

- 1) Flight number for flight i , PF_i .
- 2) Planned departure time for flight i , PD_i .
- 3) Type of aircraft.
- 4) Duration of flight i .
- 5) Number of passengers on flight i .
- 6) Departure airport of flight i .
- 7) Arrival airport of flight i .
- 8) Latest time for operations on the airport of arrival of flight i .
- 9) Actual departure of flight i , RD_i .
- 10) Actual arrival of flight i , RA_i .

The authors assume that a aircraft takes off immediately for the next flight after it has landed for a preceding flight.

$$RA_i = RD_{i+1} \quad (198)$$

Also,

$$RD_i = RA_{i+1} \quad (199)$$

Since the actual departure time cannot be greater than planned departure time

$$RD_i \geq PD_i \quad (200)$$

Let CT_i be the latest time for operations on the airport of arrival of flight i

$$RD_i + ET_i \leq CT_i \quad (201)$$

$$PD_i \leq RD_i \leq CT_i - ET_i \quad (202)$$

Delay DT_i on flight PF_i equals

$$DT_i = RD_i - PD_i \text{ if } PD_i \leq RD_i \leq CT_i - ET_i$$

$= 0$ otherwise
 $X_i = 1$ if flight i is flown
 $= 0$ otherwise
 Total time losses on the network,

$$TL = \sum_i X_i . LT_i \quad (203)$$

Total number of flight flown,

$$TC = \sum_i X_i \quad (204)$$

The problem is

$$\text{Min } TL = \sum_i X_i . LT_i$$

$$\text{Max } TC = \sum_i X_i$$

subject to:

$$X_i = 1, \forall i : PD_i \leq RD_i \leq CT_i - ET_i \quad (205)$$

$$X_i \in \{0,1\}, \forall i \quad (206)$$

Stojkovic, Soumis and Desrosiers (2002) develop a simple linear programming based model, to generate new schedules following schedule perturbations. The authors preserve flight and crew itineraries, but only change arrival and departure times.

The variables used are:

O : The set of origin nodes.

D : The set of destination nodes.

F : The set of arcs for flight legs.

G : The set of arcs for ground legs.

M : The set of arcs for maintenance.

C : The set of arcs for passenger connections.

T : The set of arcs for crew connections.

L : The set of arcs for limiting length of itineraries.

$t_i(a_i \leq t_i \leq b_i)$: Time associated with arrival and departure.

$t_{ij}(a_{ij} \leq t_{ij} \leq b_{ij})$: The duration of each flight leg.

C_i : The cost for altering the arrival and departure times.

C_{ij} : The cost for altering the flight arc duration.

The model formulation is:

Min

$$Z = \sum_i C_i t_i + \sum_{ij} C_{ij} t_{ij} \quad (207)$$

subject to:

$$t_j \geq t_i + t_{ij}, \forall (i, j) \in F \cup G \cup M \cup C \quad (208)$$

$$t_j \geq t_i + a_{ij}, \forall (i, j) \in T \cup R \quad (209)$$

$$t_j \geq t_i - b_{ij}, \forall (i, j) \in L \quad (210)$$

$$a_i \leq t_i \leq b_i, \forall i \in O \cup D \quad (211)$$

$$a_{ij} \leq t_{ij} \leq b_{ij}, \forall (i, j) \in F \cup G \cup M \cup C \quad (212)$$

Lohatepanont and Barnhart (2002) and (2004) solve the schedule generation problem in two stages:

- 1) Frequency planning
- 2) Timetable development

The model considers supply-demand interactions and integrates schedule planning and fleet assignment. The schedule design is incremental, that is the new schedule is built from a base schedule.

The model partitions the schedule into two types of flights:

- 1) Mandatory flights - Flights that must be operated
- 2) Optional flights - Flights that can be added or deleted

The variables used in the model are:

P : The set of itineraries in a market.

A : The set of airports.

L : The set of flight legs.

K : The set of fleet types indexed by k .

T : The sorted set of all events times at all airports, indexed by t_j .

N : The set of nodes.

$CL(k)$: The set of flight legs that pass the count time, when flown by the fleet type k .

$I(k, o, t)$: The set of inbound flight legs to the node $\{k, o, t_j\}$.

$O(k, o, t)$: The set of outbound flight legs to the node $\{k, o, t_j\}$.

t_p^r = the number of passengers requesting itinerary p but are redirected by the model to itinerary r .

$f_{k,i}$ = 1 if fleet leg i is assigned to fleet type k .

$y_{k,o,t}^+$ = no of aircraft on the ground at airport o , immediately after time t .

$y_{k,o,t}^-$ = no of aircraft on the ground at airport o , immediately after before t .

Parameters

CAP_i = The number of seats available on leg i .

$SEATS_k$ = The number of seats available on the aircraft in type k .

N_k : The number of aircraft in type k .

D_p : The unconstrained demand for itinerary p .

$fare_p$: The fare for itinerary p .

$fare_p'$: cost-adjusted fare.

b_p^r : recapture rate from itinerary p to itinerary r .

$\delta_i^p = 1$ if itinerary p includes flight leg i
 $= 0$ otherwise

Δ_p^q : The demand correction term for itinerary p , as a result of cancelling itinerary q .

$Z_q = 1$ if itinerary q is operated
 $= 0$ otherwise.

The authors assume that

- 1) Independent markets, that is markets do not interact with each other.
- 2) The model does not take into account higher order effects of deleted itineraries, i.e the effect of two simultaneous deletions is not the same as two consecutive deletions

The model formulation is:

Min

$$Z = \sum_{i \in L} \sum_{k \in K} C_{k,i} f_{k,i} + \sum_{p \in P} \sum_{r \in R} (fare_p - b_p^r fare_r) t_p^r + \sum_{q \in P_o} (fare_q D_q - \sum_{p \in P, p \neq q} fare_p \Delta D_q^p) (1 - Z_q) \quad (213)$$

subject to:

$$\sum_{k \in K} f_{k,i} = 1, \forall i \in L^F \quad (214)$$

$$\sum_{k \in K} f_{k,i} \leq 1, \forall i \in L^O \quad (215)$$

$$y_{k,o,t^-} + \sum_{i \in I(k,o,t)} f_{k,i} - y_{k,o,t^+} - \sum_{i \in O(k,o,t)} f_{k,i} = 0, \forall \{k,o,t\} \in N \quad (216)$$

$$\sum_{o \in A} y_{k,o,t_m} + \sum_{i \in CL(k)} f_{k,i} \leq N_k, \forall k \in K \quad (217)$$

$$\sum_{p \in P} \sum_{q \in P_o} \delta_i^p \Delta D_q^p (1 - Z_q) + \sum_{k \in K} CAP^k f_{k,i} + \sum_{r \in P} \sum_{p \in P} \delta_i^p t_p^r - \sum_{r \in P} \sum_{p \in P} \delta_i^p b_r^p t_r^p \geq Q_i, \forall i \in L \quad (218)$$

$$\sum_{r \in P} t_p^r + \sum_{q \in P_o} \Delta D_q^p \delta_i^p (1 - Z_q) \leq D_p, \forall p \in P \quad (219)$$

$$Z_q - \sum_{k \in K} f_{k,i} \leq 0, \forall i \in L(q) \quad (220)$$

$$f_{k,i} \in \{0,1\}, \forall k \in K, \forall i \in L \quad (221)$$

$$Z_q \in \{0,1\}, \forall q \in P^O \quad (222)$$

$$y_{k,o,t} \geq 0, \forall \{k,o,t\} \in N \quad (223)$$

$$t_p^r \geq 0, \forall p, r \in P \quad (224)$$

Sherali et.al (2009) developed an integrated schedule planning and fleet assignment problem that considers optional flights with existing mandatory flights along with itinerary based demands, and performed a polyhedral analysis to enhance problem solvability. This approach was exhibited to yield a significant increase in annual profits using real data from United Airlines.

2.2.3) Supply-Demand Based Formulations

Formulations presented in Sections 2.1.3 and 2.1.4 solve the Air Traffic Network Flow problem and the Schedule Generator problem simultaneously. Their strengths and drawbacks have been discussed in Sections 2.1.3 and 2.1.4.

2.2.4) Heuristic Based Formulations

Heuristic algorithms use simple rules to determine the efficiency of a schedule, and are largely driven by trial-and-error. They are based on observations of real schedules and constructing simple rules to construct new or improved schedules, without incorporating much mathematical rigor.

Gosling and Hansen (1995) test the common assumption of aircraft size being positively correlated with increased delay. The FAA's forecasts predict that average aircraft sizes across the National Airspace System (NAS) would increase from 141 to 149 from 1990 to 2000 thus resulting in a decrease in flight operations and mitigating congestion.

The authors discuss several ways of overcoming congestion:

- 1) New flights could be added at less busy times.
- 2) Service could be offered at other (secondary) airports.
- 3) Connecting passengers could be routed through other airports.
- 4) Average aircraft size could be increased.

The authors point out that increase in traffic has outpaced increase in aircraft size. At LAX traffic has tripled from 1970 to 2000 yet aircraft size has only doubled. In SFO and SEA the average aircraft size has increased only 2% whereas the increase in traffic has been 36%. The LAX-SEA market was initially served by four carriers, however two withdrew from the market and the aircraft sizes for the other two carriers (Alaskan and United) increased by about 4%, although traffic levels quadrupled. The authors estimate the elasticity of aircraft size with respect to some independent variables using linear regression. The independent variables are:

- 1) No of passengers (Market size)
- 2) Stage length
- 3) Concentration (Herman-Herfandahl index)
- 4) Airport delay levels

The regression equation used is

$$\ln(SPF_{ij}) = \alpha_0 + \alpha_1 + \sum_k \beta^k X_{ij}^k + \varepsilon_{ij} \quad (225)$$

Separate models are estimated for low and high-density markets, and short and long haul markets. The authors find that larger aircraft are used in high-density, long-haul markets. The model predicts that a 10% increase in traffic results in only a 2% increase in aircraft size for markets less than 500 miles, dropping to 1% for markets above 2500 miles.

The authors conclude that:

- 1) Increase in traffic volumes has not produced an appreciable increase in aircraft size.
- 2) Delays lead to modest increase in aircraft size, since market share is proportional to flight frequency.
- 3) The prospects for a larger aircraft size are not encouraging on the basis of market forces alone, therefore airport policy decisions have to be made by a regulatory body.
- 4) Further analysis of competition between airlines is needed.

Dennis (2000) compares the efficiency of Amsterdam Schipol and London Heathrow airport and concludes that for maximum connecting efficiency, it is desirable for the flights to be scheduled in four or five “wave” patterns as implemented in Schipol. On the other hand connecting times at Heathrow are longer on account of Heathrow following a uniform flight schedule density.

The problem of wave schedules vis-à-vis uniform schedules is explored in greater depth by Rietveld (2001). The author points out that higher frequencies do not necessarily mean lower transfer times for passengers. The author derives a schedule coordination factor (α) for hub airports. The author presents the schedule coordination factors for Schipol and Heathrow and concludes that Schipol has a higher coordination factor (0.8) versus 0.5 for Heathrow on account of the wave pattern of schedules at Schipol.

The advantage of the heuristic based formulations especially for large systems, is their inherent simplicity and applicability to real-world problems. However they suffer from the drawback that heuristics are not good agents to predict the future states of the system, since they assume the same system behavior in the future. Nevertheless heuristics are extensively used in network modeling and forecasting since traditional rigorous approaches are often too expensive in terms of computation time or the explanatory variables unknown which preclude any mathematical approach.

Table 2.2 Methodologies Used for the Schedule Generator Problem.

Methodology	Benefits	Drawback	Characteristics	Size of problem
Mathematical programming, independent of previous schedule	Can construct the most profitable schedule.	May not reflect the real operating schedules of the airlines.	Linear programming, Integer programming, Goal programming	A single route to the entire American Airlines network
Mathematical programming, dependent of previous schedule	Can be used to mimic the real airline schedule.	May not construct the most profitable airline schedule.	Linear programming, Integer programming, Network flow	A few flight legs to the network for a large domestic carrier
Supply-demand based formulations	Both demand and supply-side variables are modeled	Iteration for equilibrium.	Maximum likelihood	31 hub airports, Taiwan domestic air network
Heuristic based formulations	Computational simplicity	Governing rules difficult to obtain	Non mathematical rule based	N/A

2.3) Literature review on the Route Generator Problem

The objective of the Route Generator Problem is to expand the existing network by adding new routes or new hub or spoke airports. The output of the route generator problem would be the new route opened, the markets served by the new route, and the hub serving as an intermediate airport for the new route, if the route is a non-direct route. The Route Generator Problem can be formally described formally as:

Given a base network, the demand for each market in the network, and the fleet type serving each arc in the network, and the direct and indirect costs of building or establishing a new airport

Find the set of routes to be added to the network that minimizes the congestion for the passengers and maximizes the airline profit

Various techniques have been used in literature to model the location of new hubs, airports and the construction of new routes. They can be broadly classified as:

- (1) Mathematical programming based formulations.
- (2) Econometrics based formulations.
- (3) Heuristic based formulations.

2.3.1) Mathematical Programming Based Formulations

Mathematical programming based approaches primarily attempt to find the location of a hub airport, while minimizing the cost and considering other constraints. These models assume that the present network is at or has exceeded capacity. The models primarily employ integer or mixed integer programming to locate the best hub among a set of candidates.

O'Kelly (1986) describes and solves the problem of hub locations, considering interaction effects. The author assumes that the cost of location of a hub is dependent on both its location and the location of other hubs.

The variables used are:

$D_k : (X_k, Y_k)$ be the position of the hub

$(x_i, y_i) : \text{Set of nodes(cities).}$

The model formulation is:

One-hub problem:

Min

$$\sum_i \sum_j W_{ij} * (C(p_i, Q) + C(Q, p_j)) \quad (226)$$

Where $C(p_i, Q)$ = cost of serving the node p_i from the hub Q .

Two-hub problem

Let $U_i = 1$ if node i is served by hub 1

= 0 otherwise

Let $V_i = 1$ if node i is served by hub 2.

= 0 otherwise

Min

$$\sum_i \sum_j W_{ij} (U_i U_j (C_{i1} + C_{j1}) + V_i V_j (C_{i2} + C_{j2}) + U_i V_j (C_{i1} + C_{j2} + C_{12}) + U_j V_i (C_{i2} + C_{j1} + C_{21})) \quad (227)$$

where:

$$C_{i1}, C_{i2}, C_{j1}, C_{j2} = f(X_i, Y_i) \quad (228)$$

The above model restricts the number of hubs to one or two. In some cases, especially for national carriers, it might be necessary to establish multiple hubs. Researchers therefore started to design models which could handle a p-hub problem.

O'Kelly (1987), (38) proposes a model to solve the p-hub problem, considering interaction effects. The variables used in the model are

$X_{ik} = 1$ if city i is assigned to hub k .

= 0 otherwise.

$X_{ii} = 1$ if city i is a hub.

= 0 otherwise

W_{ij} : flow between i and j . = $f(d_{ij})$.

C_{ij} : cost of service between i and j .

The model formulation is,

Min

$$\sum_j \sum_i W_{ij} (\sum_k C_{ik} X_{ik} + \sum_m C_{mj} X_{mj} + \sum_k \sum_m X_{ik} X_{jm} C_{km}) \quad (229)$$

subject to:

$$\sum_k X_{ik} = 1, \forall i \quad (230)$$

$$\sum_k X_{kk} = p \quad \text{Where } p \text{ is the number of hubs to be established} \quad (231)$$

$$(n - p + 1)X_{jj} - \sum_i X_{ij} \geq 0, \forall j \quad (232)$$

$$X_{ij} \in \{0,1\}, \forall i, j \quad (233)$$

The two proposed models are quadratic integer programs, similar in structure to the quadratic assignment problem (QAP), which is very difficult to solve for any problem greater than 15 nodes. The author's formulation also disregards economies of scale arising from hub networks.

Aykin (1995) proposes a linear integer program for hub location. In contrast to O'Kelly's formulation, the author's formulation focuses on the profitability of a route rather than a hub. The author also considers two hubbing policies in an attempt to mimic airline behavior:

- 1) Nonstrict hubbing in which channeling flows through hubs is not required but chosen if found cost efficient.
- 2) Strict and restrictive hubbing in which all flows to/from a node are channeled through a hub.

The variables used are:

$Y_k = 1$ if a hub is located at k .

$= 0$ otherwise

$X_{ij} = 1$ if a direct flight is provided between i and j .

$= 0$ otherwise

$X_{ikj} = 1$ if hub connected service is provided from node.
 i to node j with the routing $i \rightarrow k \rightarrow j$.

$= 0$ otherwise.

S_{ij} : cost of service between i and j .

d_{ij} : distance between i and j .

$\Gamma_{ij} = 1$ if direct flights are considered between i and j .

$= 0$ otherwise

a, a_1, a_2 : Scale economy factors.

The formulation for the nonstrict hubbing case is:

Min

$$Z = \sum_i \sum_j W_{ij} * (S_{ij} d_{ij} X_{ij}) + \sum_i \sum_k \sum_t \sum_j W_{ij} * (a_1 S_{ik} d_{ik} + a S_{kt} d_{kt} + a_2 S_{tj} d_{tj}) * X_{iktj} + \sum_k F_k * Y_k \quad (234)$$

subject to:

$$X_{ij} + \sum_k \sum_t X_{iktj} = 1, \forall i, j \quad (235)$$

$$\sum_k Y_k = p \quad (236)$$

$$\sum_{i \neq k} \sum_{l \neq k} (X_{ilhk} + X_{klhi}) + \sum_i (X_{ik} + X_{ki}) \leq M(1 - Y_k), \forall k \quad (237)$$

$$\sum_i \sum_j \sum_t (X_{iktj} + X_{itkj}) \leq MY_k, \forall k \quad (238)$$

$$\sum_{h \neq k \neq t} (X_{kkht} + X_{thkk} + X_{tthk} + X_{khtt}) + \sum_h (X_{khht} + X_{thhk}) \leq M(2 - Y_k - Y_t) \quad (239)$$

The formulation for the strict restrictive hubbing case is:

$$Z = \sum_i \sum_j W_{ij} * \Gamma_{ij} * (S_{ij} d_{ij} X_{ij}) + \sum_i \sum_k \sum_t \sum_j W_{ij} * (a_1 S_{ik} d_{ik} + a S_{kt} d_{kt} + a_2 S_{tj} d_{tj}) * X_{iktj} + \sum_k F_k * Y_k \quad (240)$$

$$\Gamma_{ij} X_{ij} + \sum_k \sum_t X_{iktj} = 1, \forall i, j \quad (241)$$

$$\sum_k Y_k = p \quad (242)$$

$$\sum_{i \neq k} \sum_{l \neq k} (X_{ilhk} + X_{klhi}) + \sum_i (X_{ik} + X_{ki}) \leq M(1 - Y_k), \forall k \quad (243)$$

$$\sum_i \sum_j \sum_t (X_{iktj} + X_{itkj}) \leq MY_k, \forall k \quad (244)$$

$$\sum_{h \neq k \neq t} (X_{kkht} + X_{thkk} + X_{tthk} + X_{khtt}) + \sum_h (X_{khht} + X_{thhk}) \leq M(2 - Y_k - Y_t) \quad (245)$$

However, the above formulations assume a predefined network structure and do not consider aircraft costs in the formulation. Jalliet and Song (1996) propose a model that does not assume a network structure, and explicitly accounts for aircraft operating costs. The variables used in the model are:

N : set of all cities

$d_{ij} = d_{ji}$: distance between cities i and j .

$f_{ij} = f_{ji}$: No of people that desire to travel between i and j .

K : set of different types of aircraft to choose from.

C_k : Cost/mile.

B_k : Capacity.

X_{ijl} : Fraction of the passengers desiring to travel from i to j , through l .

$O(d)$: Origin node for city pair $d \in D$.

$D(d)$: Destination node for city pair $d \in D$.

f_d : Demand for the O-D pair $d \in D$.

Separate models are formulated for one-stop, two-stop and all stop.

The formulation for the one-stop model is:

Min

$$\sum_{i \neq j} \sum_{k \in K} d_{ij} c_k y_{ij}^k \quad (246)$$

subject to:

$$f_{ij} + \sum_{t \neq i, j} (f_{it} x_{ijt} + f_{tj} x_{tij} - f_{ij} x_{itj}) \leq \sum_{k \in K} b_k y_{ij}^k, \forall i \neq j \quad (247)$$

$$\sum_{t \neq i, j} x_{itj} = 1, \forall i \neq j \quad (248)$$

$$x_{itj} \geq 0, \forall i \neq t \neq j \quad (249)$$

$$y_{ij}^k \geq 0 \text{ and integer}, \forall i \neq j \quad (250)$$

The formulation of the two-stop model is

Min

$$\sum_{i \neq j} \sum_{k \in K} d_{ij} c_k y_{ij}^k \quad (251)$$

subject to:

$$\begin{aligned} & f_{ij} + \sum_{t \neq i, j} (f_{it} x_{ijt} + f_{tj} x_{tij} - f_{ij} x_{itj}) + \\ & \sum_{l, t \neq i, j} (f_{lj} x_{litj} + f_{it} x_{ijlt} + f_{lt} x_{lijt} - f_{ij} x_{iltj}) \\ & \leq \sum_{k \in K} b_k y_{ij}^k, \forall i \neq j \end{aligned} \quad (252)$$

$$\sum_{t \neq i, j} x_{itj} + \sum_{l, t \neq i, j} x_{iltj} = 1, \forall i \neq j \quad (253)$$

$$x_{itj} \geq 0, \forall i \neq t \neq j \quad (254)$$

$$x_{iltj} \geq 0, \forall i \neq l \neq t \neq j \quad (255)$$

$$y_{ij}^k \geq 0 \text{ and integer}, \forall i \neq j \quad (256)$$

The formulation of the all stop model is

Min

$$\sum_{i \neq j} \sum_{k \in K} d_{ij} c_k y_{ij}^k \quad (257)$$

subject to:

$$\sum_{j \neq i} z_{ij}^d - \sum_{j \neq i} z_{ji}^d = \begin{cases} f_d \rightarrow i = D(d) \\ -f_d \rightarrow i = O(d) \\ 0 \rightarrow otherwise \end{cases} \quad (258)$$

$$\sum_{d \in D} z_{ij}^d \leq \sum_{k \in K} b_k y_{ij}^k, \forall i \neq j \quad (259)$$

$$z_{ij}^d \geq 0, \forall i \neq j, k \in K \quad (260)$$

$$y_{ij}^k \geq 0 \text{ and integer}, \forall i \neq j \quad (261)$$

The authors solve the above MIP's by adopting a simulated-annealing based greedy interchange algorithm.

As airlines added more airports to their networks, competition between airlines often on the same routes or two or more airlines sharing the same hub or route has become increasingly common. The models described above optimize the hub location, considering the payoff only when one player is present. Clearly if there is competition, then the hub locations derived from the above models may not be correct.

Marianov, Serra and ReVelle (1999) present a model to solve the hub location while considering competitive effects.

The variables used are:

Formulation I

$w_{kl} = 1$ if hubs are located at k and l .

$= 0$ otherwise

p - The number of hubs to be established.

$x_k = 1$ if a hub is located at k .

$= 0$ otherwise

a_{ij} = flow from i to j .

c_{ij} = current cost for carrying a unit of traffic from i to j .

C_{ij} = competitor's cost for carrying a unit of traffic from i to j .

y_{ij} = proportion of the flow captured between i and j .

Formulation II

r_{ij} = revenue per unit flow between i and j .

f_k = fixed cost of constructing a hub at k .

Formulation III

$y^{50} = 1$ if 50% of the flow between i and j is captured.

$= 0$ otherwise

$y^{75} = 1$ if 75% of the flow between i and j is captured.

$= 0$ otherwise

$y^{100} = 1$ if 100% of the flow between i and j is captured.
 $= 0$ otherwise

$$N_{ij}^{50} = \{(k,l) | 0.9C_{ij} \leq C_{ik} + \alpha C_{kl} + C_{lj} \leq 1.1C_{ij}\}.$$

$$N_{ij}^{75} = \{(k,l) | 0.7C_{ij} \leq C_{ik} + \alpha C_{kl} + C_{lj} \leq 0.9C_{ij}\}.$$

$$N_{ij}^{100} = \{(k,l) | 0 \leq C_{ik} + \alpha C_{kl} + C_{lj} \leq 0.7C_{ij}\}.$$

Three formulations are given, the first formulation considers minimization of the passenger cost, the second considers the fixed cost of hub location, in addition to the passenger cost and the third formulation proposes that the route flow is not an “all-or-nothing flow”, i.e. passengers can also be routed along less profitable routes. The three formulations are

Formulation I

Max

$$Z = \sum_{i \in I} \sum_{j \in J} a_{ij} y_{ij} \quad (262)$$

subject to:

$$w_{kl} \leq x_k \quad (263)$$

$$y_{ij} \leq \sum_{(k,l) \in N} w_{kl} \quad (264)$$

$$\sum_{k \in K} x_k = p \quad (265)$$

Formulation II

Max

$$Z = \sum_{i \in I} \sum_{j \in J} r_{ij} y_{ij} - \sum_{k \in K} f_k x_k \quad (266)$$

subject to:

$$w_{kl} \leq x_k \quad (267)$$

$$w_{kl} \leq x_l \quad (268)$$

$$y_{ij} \leq \sum_{(k,l) \in N} w_{kl} \quad (269)$$

$$\sum_{k \in K} x_k = p \quad (270)$$

Formulation III

Max

$$Z = \sum_{i \in I} \sum_{j \in J} \alpha_{ij} y_{ij}^{100} + 0.75 \alpha_{ij} y_{ij}^{75} + 0.5 \alpha_{ij} y_{ij}^{50} \quad (271)$$

subject to:

$$w_{kl} \leq x_k, w_{kl} \leq x_l \quad (272)$$

$$y_{ij}^{100} \leq \sum_{(k,l) \in N_{ij}^{100}} w_{kl}, y_{ij}^{75} \leq \sum_{(k,l) \in N_{ij}^{75}} w_{kl}, y_{ij}^{50} \leq \sum_{(k,l) \in N_{ij}^{50}} w_{kl}, \quad (273)$$

$$y_{ij}^{100} + y_{ij}^{75} + y_{ij}^{50} \leq 1, \sum_{k \in K} x_k = p \quad (274)$$

Lederer and Nambimadom (1998) propose a model that considers various network configurations, and derives the optimal passenger demand for each of the networks.

Models that employ mathematical programming are rigorous and make few simplifying assumptions. However these mathematical formulations are binary integer programs which might be difficult to solve for large networks.

2.3.2) Econometric Based Formulations

Econometric based models derive the probability of a city being a hub based on regression or maximum likelihood approaches. The models need a set of input cities, whose viability of being a hub are derived.

Bauer (1987) proposes a model to estimate the probability of a city being chosen a hub. The variables chosen are:

pop_i = Population of city i .

inc_i = Income of city i (local demand).

$DBTP_i$ = Dummy variable for business/recreation.

$corp_i$ = No of fortune 500 corporations in the city.

rec_i = Rank of the city i in recreation.

$cult_i$ = Rank of the city i in culture.

The model formulation is:

$$h_i = a_0 + a_1 \ln(pop_i) + a_2 \ln(inc_i) + a_3 \ln(DBTP_i) + a_4 \ln(corp_i) + a_5 * \ln(rec_i) + a_6 \ln(cult_i) + v_i$$

Where

$h_i = 1$ if hub is viable at location i

$= 0$ otherwise

The author assumes that if $h_i \geq k$ (threshold value), $h_i = 1$. The model is estimated using a Probit maximum likelihood estimate.

2.3.3) Logit Based Formulations

Logit based formulations employ socio-economic variables to obtain the profitability of a network configuration. Given the increasing delays at large hubs in recent years, the

principal benefits of a hub-and-spoke network are being eroded. Some airlines therefore have started establishing point-to-point networks.

Kanafani and Ghobrial (1984) examine the reasons for hubbing (economies of scale, increased flight frequencies etc.) and the historical evolution of the hubbing phenomenon. The authors conclude that the implications of the hubbing phenomenon on airport economics are three fold:

- 1) The revenue generated by the connecting passengers is less than that generated from originating passengers.
- 2) The airport is dependent on the airline network.
- 3) The hub airports do not contribute to the local economy.

The study uses a multinomial logit model to study route choice, then an aircraft and delay model to evaluate the delay to the airport. The authors hypothesize that if an airport charges landing fees due to increased operations, then the airline reacts by scaling back frequency or increasing airfare. The model studies the effect of congestion/landing fees by airports on airlines that hub.

The authors conclude that:

- 1) The landing fees or congestion accrued by the carriers due to hubbing has not proven to be a significant factor in the hubbing phenomenon since the airlines pass the costs directly to the passengers either in the form of reduced frequencies or increased air fares.
- 2) Hubbing represents a good opportunity for the airlines to earn money.
- 3) In the future the present hubs will reach capacity and become unattractive. Therefore new hubs will emerge. Hubbing is going to be predominant for the foreseeable future.

The advantage of logit models in deciding route structure, are their computational simplicity vis-à-vis mathematical programming. However logit models assume that every cross-section of travelers behaves in the same way, and therefore might not be adequate to capture the variation across the population.

2.3.4) Heuristic Based Formulations

Simpson (1969) outlines a model for route extension on a linear point-to-point network.

Four basic methods of route planning are considered:

- 1) Create a new route.
 - 2) Add an extension to an existing route.
 - 3) Insert a city in an existing route.
 - 4) Join or bridge together two new routes.
- 1) Creation of a new route.

Three types of routes are considered - zero stop, one stop and two stop. New routes are considered whenever demand justifies a new flight, and flight frequencies are determined from aircraft size and break-even load factor.

- 2) Adding an extension to an existing route.

Extension can be done by:

- (i) Extending a non-stop to a one-stop route.
 - (ii) Extending a non-stop twice to a two-stop.
 - (iii) Adding two ends to a non-stop to form a two-stop.
- 3) Insertions of a city into a new route.
The authors discuss three types of insertions one city, two cities and three cities to form a three-stop, four stop and five stop route
Authors also point out that the correct way of calculating the load factor
- 4) Bridge together two routes.
The author presents several ways to bridge together two routes and warn that the frequencies of the combined route may not be the same as that of the individual route.

The basic plan given is:

- 1) Begin with the set of all traffic demands into routes.
- 2) Select the aircraft types.
- 3) Generate all possible zero-stop routes, calculate unmet demand.
- 4) Generate all one-stop routes, calculate unmet demand.
- 5) Generate all two-stop routes, calculate unmet demand.
- 6) Repeat for all three and four stop routes.
- 7) Investigate route joining and bridging.
- 8) Investigate route extensions.
- 9) Investigate route insertions.
- 10) Stop whenever the fleet for the aircraft type is exhausted.
- 11) Repeat with another aircraft type.

Veldhuis (1997) derives an attractiveness index for the network in Western Europe based on the quality of connections

The author presents a method to determine the quality of connections at a hub. The author groups the airports into:

- 1) Main ports in Western Europe.
- 2) Secondary airports in Western Europe.
- 3) Other airports in Western Europe.
- 4) Other airports in Europe outside Western Europe.
- 5) Other airports outside Europe.

Attractiveness = $f(\text{Travel Time, Frequency, Fares})$

No of connectivity units from/to the airport = $\text{QUAL} * \text{FREQ}$

QUAL = Quality index

FREQ = Frequency

$\text{QUAL} = 1 - (\text{PTT} - \text{NST}) / (\text{MAXT} - \text{NST})$

$\text{MAXT} = (3 - 0.075 * \text{NST}) * \text{NST}$

$\text{PTT} = \text{FLY} + 3 * \text{TRF}$

Where

MAXT = Maximum perceived Travel Time

NST = Non-stop Travel Time

FLY = Flying Time

TRF = Transfer Time

CNU = No of connectivity units

The author applies this model to the connectivity performance of Amsterdam Schipol.

Weber (2001) examines the reasons for long-haul route development. Five distinct perspectives: geography, regulation, manufacturers, passengers and airlines are taken to investigate the drivers of long-haul route development. From the insights gained it is inferred that the demand for long-haul aircraft is a complex function thereof, whereby long-haul routes are operated between cities with strong business and social connections as influenced by the dynamics of change and underlying geography.

Swan (2002) describes the historical trends in route developments for 1985-2000. The author points out that the intuitive thinking that airlines enhance service on existing routes first, before adding new routes is not true as growth in new routes has been the same as enhancements to existing routes. Also growth in small airplanes has been substantial. The airlines have responded to travel demand by more links/routes and smaller aircraft types rather than bigger airplanes. Growth is absorbed by more frequencies, routes, airports and not larger airplanes. The author argues that deregulation had led to development of new routes and smaller airplanes. Also economies of scale tend to diminish as airplanes get larger, and adding new routes saves costs by evading competition. The author points out that congestion have made the carriers cut the small airplanes (19 seats) although the average aircraft size has diminished.

Savage and Scott (2004), (39) estimate a multiple linear regression model to predict the spoke cities, that are likely to be connected to a hub. The model is used to predict the 36 new links that were added by Delta Air Lines affiliate Comair to its Cincinnati/Northern Kentucky Airport hub between 1996 and 2001. The independent variables considered are:

- 1) Distance
- 2) Population of the Metropolitan Statistical Area at the end of the spoke
- 3) Demand between the hub and the spoke
- 4) Ratio of the distance between the spoke and hub, via a secondary airport to the direct spoke-hub distance
- 5) No of flights from the spoke to the secondary airport.
- 6) The existence of a city within 60 miles of the spoke, having a direct service to a hub

The authors conclude that the probability of new service to Cincinnati drops significantly with distance, suggesting that RJs are mainly used for short haul routes. The effect of O-D demand on the frequency of RJ service to Cincinnati suggests that the RJs are used to attract itinerant passengers rather than direct passengers. The existence of a secondary airport with direct service was not found to be statistically significant for estimation of the frequency. The number of flights to the secondary airport was positively correlated to the frequency, which suggested that already existing service is an incentive for market entry.

Lee and Ito (2005) presents a trend of code-sharing practices in the US airline industry and explain that code-sharing is common in the international air travel market since no

single airline can serve all the markets, but the effect of code-sharing on the domestic airline industry has been limited, with 98% of the passengers served by the same airline. Nevertheless code-sharing in the domestic airline industry has been increasing in recent years. The authors point out that there are three advantages of code-sharing:

- 1) Airlines can pursue network expansion by capturing key markets.
- 2) Code-sharing avoids double marginalization which occurs between non-allied interline carriers.
- 3) Code-sharing could be used for product differentiation, leading to lower fares.

Gillen and Morrison (2005) explore the reasons for the operational efficiency and profitability of the low-cost point-to-point carriers or VBAs (Value Based Airlines)

The authors explain that the profitability of an airline is dependent on the aircraft operating paradigm, and the effects of supply-demand economics

On the supply side, a hub-and-spoke network creates economics of density and a point-to-point network creates an aircraft utilization density. Therefore, a hub-and-spoke network is only profitable when economics of density are greater than aircraft utilization economics. On the demand side, the network effects of a hub-and-spoke system create the need for a complex revenue-management system that focuses on the profitability of the whole network. On the other hand the simple revenue management system of a point-to-point network maximized the revenue of each link independently of the other links.

The operating principle of the full-service carriers leads to high labor and operating costs, complex schedules and dynamic yield management. Full-service airlines maximize revenue rather than cost, while the low cost carriers operate a homogenous fleet, have simple flight schedules and use secondary airports. The operating principle is aimed to reducing cost rather than maximizing revenue.

The authors analyze two cases

- 1) Fully connected network
- 2) Hub-and-spoke network

Hub-and-spoke networks are more profitable when the operating costs are high. The fully connected network is preferred network structure when the fixed costs and returns to density are low

The authors conclude by outlining strategies that the full service airlines have used against their low-cost competitors

- 1) Adoption of some of the operating paradigms of the low-cost carriers (reduced in-flight service, more direct service and operation of low-cost subsidiaries.
- 2) Simplification of organization rules.
- 3) Promotion of code-sharing to expand network size.

Lee (2005) analyzes concentration and price changes in the US domestic airline industry throughout the 1990s. In general, the authors find that national measures of concentration remained relatively flat throughout the 1990s while overall airline prices fell significantly in real terms. These findings tend to suggest that concerns of increasing concentration and rising prices are largely unfounded.

Table 2.3 Methodologies Used for the Route Generator Problem.

Methodology	Benefits	Drawbacks	Characteristics	Size of the problem
Mathematical programming	Mathematically rigorous	May be difficult to solve for large networks	Integer programming, Mixed integer programming	10 city problem to 40-city problem,
Logit based models	Computationally simple	Requires knowledge of network topology	Maximum likelihood	25 city-network, one-hub
Econometric models	Computationally simple	May not have good R^2 .	Probit regression	115 city network
Heuristic based methods	Computationally simple	Simple, rule-of-thumb rules	Forming generic rules difficult	N/A

Chapter 3 Methodology

3.1) Introduction

The transportation system analysis model (TSAM) is a nationwide multimode decision support system developed to derive the viability, cost, ridership and other service parameters for Commercial Air (CA), Automobile and the proposed Air Taxi Service system. TSAM computes the demand for each of the above modes by employing a logit framework. In order to compute the demand for CA, TSAM requires service parameters such as travel cost and travel time. Modeling the behavior of Commercial Aviation networks is, therefore key to realistically predict CA demand in the future. The Demand Based Airline and Airport Evolution model would also enable the modeler to derive the various possible states of the NAS, and suggest possible strategies to alleviate the problems facing the system today.

3.2) The Demand Based-Airline and Airport Evolution

As mentioned in Section 1.2 the Demand Based Airline and Airport Evolution Study, focuses on predicting the schedules, fares and network structures of the airline industry every year from 2000 to 2030. This can be accomplished in three ways

- (1) Treat all the airlines as a single agent or player and extrapolate their behavior into the future assuming all the carriers, will behave in the same way, ignoring competitive effects
- (2) Treat the lowcost airlines (Southwest, Jetblue) separately from the mainline carriers (United, American) since they have different operating and cost paradigms
- (3) Treat every airline as a player trying to maximize their benefit, in a non-cooperative fashion

The first approach has been widely studied in literature (See Weidner (1995), Hansen (2004) for details). This approach while, the simplest to adopt and execute is not very desirable for predicting the future state or states of the system, since interaction between various agents in the system (airlines, airports) have a significant effect on system evolution. Also the first approach cannot model collusion (code-sharing, mergers) between players, and its effect on the system. Nevertheless the simplicity of this approach makes it attractive for large networks.

The second approach takes into account the assault of the low-cost carriers on the legacy airlines, and the eroding market share and profits of the legacy carriers. In this approach the demands on each route in the air network are derived by a logit model, and in this step no distinction is made between the regular and the low-cost carriers. In the schedule generation and route generator phase, separate models are solved for the legacy and low-cost carriers, since their networks and cost parameters are different. The output of the

second step is taken as a feedback to the first step and the process repeated until equilibrium.

The third-approach is a game-theoretic approach, where each airline (lowcost and legacy) try to maximize their own profit in a non-cooperative fashion. For example a carrier might sacrifice market share in an O-D pair, where it is suffering a loss, for capturing passengers in a more profitable market. US Airways closed its services in Pittsburgh, to cut costs and to focus on more profitable markets. These effects cannot be captured without modeling the operations of each individual airline. Researchers have proposed models that try to capture the operations of an individual airline. The most notable efforts are by Hansen (1990), Hong and Parker (1992), Dobson and Lederer (1993) and Adler (2001), (2005). However these authors only consider small networks and a handful of competing carriers. However the NAS has 443 Commercial Airports and 53 individual airlines in the year (2000). The game-theoretic models are mixed-integer non-linear programs (MINLP), which are a class of difficult global nonconvex optimization problems and are generally NP-complete.

Therefore, the most desirable method for the Based Airline and Airport Evolution Study has to take into account model detail and computational tractability. The second and third approaches also require that the model predict the number of airlines into the future, which a daunting task is given the current state of the economy. Therefore, the first approach is adopted based on model detail and computational tractability.

3.3) A Feedforward Based Approach to the Demand Based Airline and Airport Evolution.

In Section, 3.2 it has been stated that the demand based airline and airport evolution problem will be solved considering that all the airlines behave as a single agent. The low cost and traditional carriers have not been considered as separate agents, since the cost differential between the low-cost and traditional carriers have declined sharply in recent years. Figure 3.1 illustrates that the full service carriers have cut their expenses sharply from 2000 to 2003 to achieve profitability. Therefore, it could be argued that the operating cost per seat-mile for the low-cost and full service carriers would eventually converge to a single value. The full-service and low-cost carriers could be treated as a single agent without any appreciable loss of model fidelity.

The demand based airline and airport evolution is composed of four sub-modules

- (i) The Air Traffic Flow Module
- (ii) The Schedule Generator
- (iii) The Fare Prediction Module
- (iv) The Network Evolution Module

These modules are discussed in detail in the subsequent sections. The entire process is shown as a flowchart in Figure 3.2

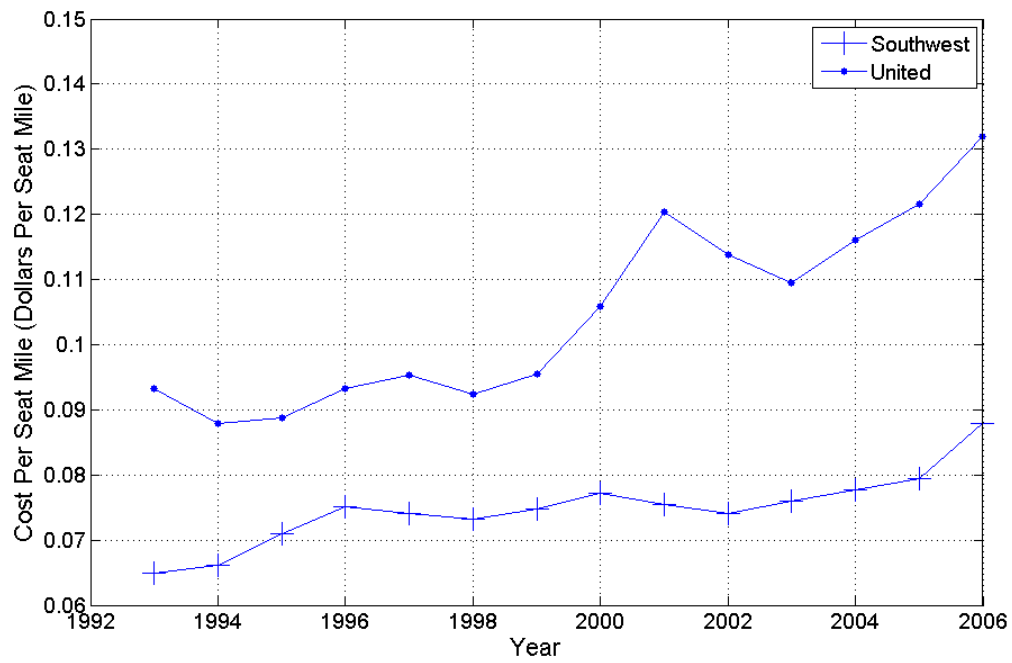


Figure 3.1 Cost per Available Seat-mile for United and Southwest Airlines (cents per ASM) (Form 10-K, Securities and Exchange Commission).

3.3.1) The Air Traffic (O-T-D) Flow Module

The Air Traffic Flow Problem can be described as:

Given a set of Origin-Destination pairs (airports), the demand between them, the set of paths covering each O-D pair and their service attributes (travel time, fares, frequency, seats) and the set of flight legs in the network. Find the demand for each flight leg, subject to link and node capacity and other additional constraints

The Air Traffic Flow (O-T-D) problem is composed of

- (i) A schedule building module for the base year, since the schedule generator is not executed for the base year
- (ii) An itinerary building module that derives all feasible itineraries between any two O-D pairs
- (iii) A route choice module that derives the probability of choosing each route using a logit framework.

The schedule-building module derived a representative yearly schedule from the Official Airline Guide (OAG). The airline schedule is not constant for the entire year as the air carriers vary the arrival and departure times to meet seasonal variations in demand. Therefore, an effective yearly schedule was constructed by taking the average of the arrival and departure times of flights that were within 30 minutes of each other. It was assumed that an airline would not schedule flights between the same Origin-Destination pair within 30 minutes of each other. Any flight, which operated less than 250 times in a year (5 days per week) was removed.

The Network Building Module, as described in the next section was constructed to obviate the need to derive the network topology from the DB1B. This step is necessary since in future years TSAM would not have access to the DB1B database. Another reason is that the DB1B yields poor samples for thin markets and a network topology based on the DB1B would not be reliable.

3.3.1.1) The Network Building Module

The network is built-up in stages. First, the O-D pairs that can be connected by direct flights are connected, followed by O-D pairs with two, three and more than three legs.

One Leg: The records that had direct flights from an origin to a destination airport are extracted from the OAG. Any record with an unusually long flying time (multi-stop flights) was removed. The average of the remaining flight times were averaged and reported as the travel time from i (Origin) to j (Destination). However, for many O-D pairs that have direct service a significant proportion of the passengers used connecting itineraries. Therefore, the two-leg solution was appended with the one-leg network for all O-D pairs having direct service.

Two Legs: The list of all the airports that have service from origin airport (i) was obtained. These O-D pairs could not be connected by a direct flight; therefore, the intermediate airports that connect them with two legs were determined. The travel time from Origin (i) to Destination (j) was the sum of the flight time from i to k , the flight time from k to j and the connecting time at airport k . The reported travel time was the average of the travel times of all feasible schedules along all the possible routes.

Three Legs: The list of all the large and medium hubs (FAA definition) ($H1$, $H2$) to which the origin and destination airport have service were obtained. It was assumed that the large and medium hubs were connected by direct flights, as about 60% of large and medium hubs are connected by direct flights in 2004. As in the previous cases, the travel time from i (Origin) to j (Destination) was the sum of the flight times from i to $H1$, $H1$ to $H2$, $H2$ to j and the connecting times at $H1$ and $H2$. Large connecting times (large travel times) were eliminated by a rolling average method.

Multi (4 or more) Legs: There were a few Origin-Destination pairs, which were unconnected even after using three legs. The potential number of combinations for any itinerary that consisted of more than three legs was too large to be derived from first principles. For these airports, a recursive approach that utilized the previous solution was adopted. The set of intermediate airports that had service to destination airport were determined. If any of these intermediate airports had an itinerary to the origin airport, the final itinerary was the addition of another flight that connected the intermediate and destination airports to the initial itinerary. Itineraries with long travel times were eliminated by a rolling average method.

Figures 3.3 and 3.4 illustrate the sample network between Roanoke, VA (ROA) to Chicago O'Hare, IL (ORD) and Lewistown, MT (LWT) derived from the network builder.

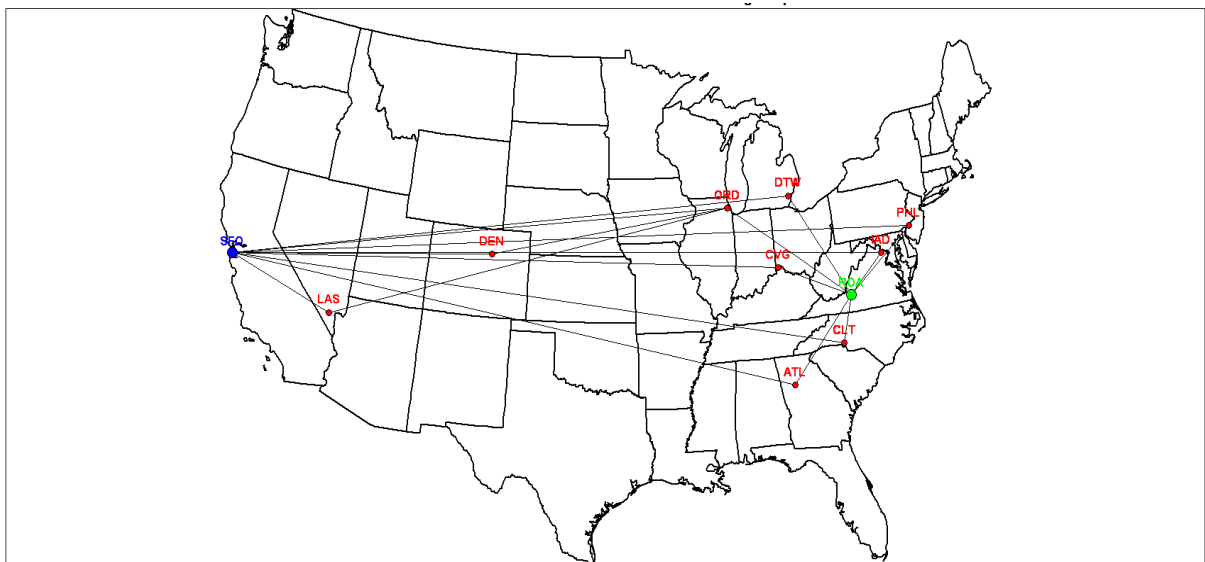


Figure 3.3 Sample Network from Roanoke, VA (ROA) to San Francisco, CA (SFO).

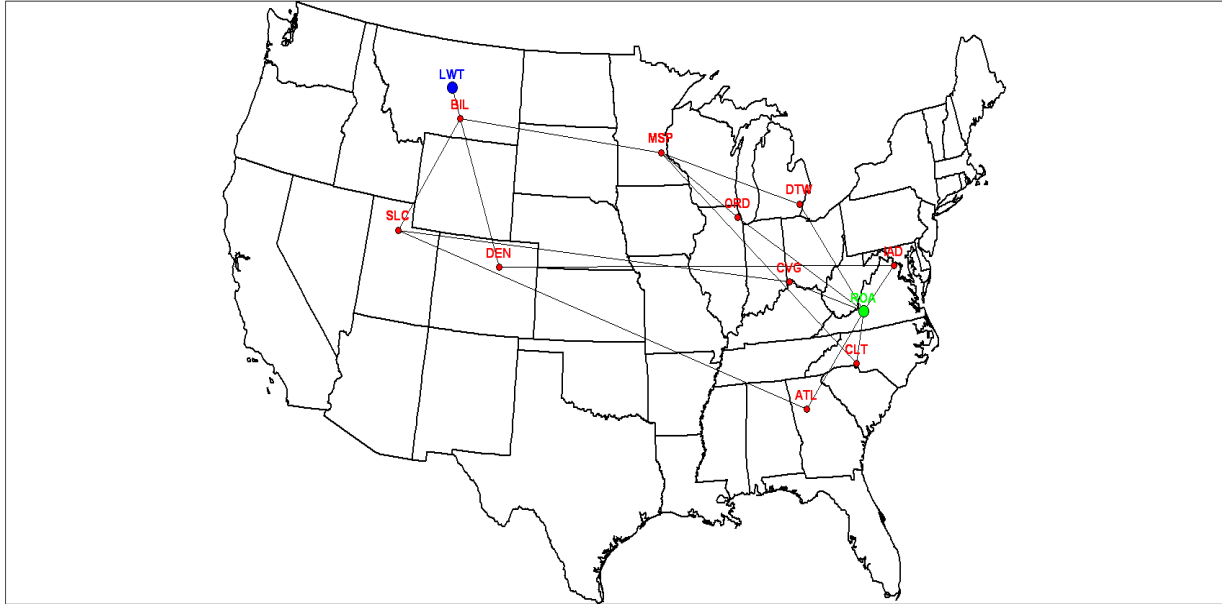


Figure 3.4 Sample Network from Roanoke, VA (ROA) to Lewistown, MT (LWT).

3.3.1.2) The Route Choice Module

The Route Choice Module used a logit model to compute the flows on each route between any two O-D pairs given the demand. The first step in the logit model was to determine the independent variables. The independent variables considered were

- (i) **Relative Travel Time:** The relative travel time was the quotient of the actual travel time on the route and the average travel time for the O-D pair. Travel times were normalized to avoid scale biases.
- (ii) **Relative Travel Cost:** The relative fare was the quotient of the actual fare on the route and the average fare for the O-D pair. As in the case of travel time, fares were normalized to avoid scale biases.
- (iii) **Number of Itineraries (Frequency):** The number of itineraries offered on a route was the number of connecting flight combinations that had reasonable connecting times. The maximum and minimum connecting times were taken as 2.5 hours and 40 minutes respectively. In case no O-D itineraries could be found with connecting times less than 2.5 hours, the maximum connecting time was taken as 3 hours greater than the minimum connecting time.
- (iv) **Seats:** The number of seats for an itinerary was taken as the number of seats offered on the first leg of the itinerary.
- (v) **Legs:** The number of legs was included to account for additional disutility incurred for connecting itineraries (with potential for missed connections or lost baggage).
- (vi) **Dummy Variables:** Dummy variables were introduced for airports, which had significant transfer disutility due to extraneous factors.

- (vii) **Ratio of volume and capacity:** The volume-to-capacity ratio was used as a surrogate for airport delay since delays are monotonic functions of the volume-to-capacity ratio. Delays are also difficult to compute at the airport level.

The data was segmented by the number of legs. Therefore, two logit models were developed, one for O-D pairs linked by both direct and connecting itineraries and the second logit model was for O-D pairs linked by connecting itineraries alone.

The utility function used in the model can be expressed as:

$$U_{ijk} = \alpha_1 * RTT_{ijk} + \alpha_2 * RTC_{ijk} + \alpha_3 * Freq_{ijk} + \alpha_4 * S_{ijk} + \alpha_5 * \eta_{ijk} + \beta_k \quad (275)$$

The relative travel time is given by,

$$RTT_{ijk} = TT_{ijk} + f\left(\frac{v}{c}\right)_k \quad (276)$$

Where:

i : The origin airport.

j : The destination airport.

k : The intermediate airport.

RTT_{ijk} : The relative route travel time with delay.

TT_{ijk} : The relative route travel time without delay.

RTC_{ijk} : The relative route fare.

$Freq_{ijk}$: The number of feasible itineraries on the route.

S_{ijk} : The number of seats offered by the airline from airport i to airport k .

η_{ijk} : The number of legs on the itinerary.

β_k : Dummy variable for some airports k .

$\left(\frac{v}{c}\right)_k$: The volume over capacity ratio for airport k .

The route passenger flow can be expressed as:

$$D_{ijk} = D_{ij} * \frac{\exp(U_{ijk})}{\sum_k \exp(U_{ijk})} \quad (277)$$

Where:

D_{ij} : The passenger demand from airport i to airport j .

D_{ijk} : The passenger demand from airport i to airport j through airport k .

The entire Air Traffic Flow Problem is shown as a flowchart in Figure 3.6

3.3.1.3) Airport Capacity Derivation

The airport capacity used in the route choice module was the annual service volume of the airport. As the passenger demand used is the yearly passenger demand. The FAA defines the annual service volume as a reasonable estimate of an airport's annual capacity. It accounts for differences in runway use, aircraft mix and weather conditions that would be encountered in a yearly time span.

The annual service volume is calculated as:

$$ASV = C_v \times D \times H \quad (278)$$

Where:

ASV : Annual Service Volume.

D : Ratio of the annual demand to the average daily demand during the peak month.

H : Ratio of the average daily demand to the average peak hour demand during the peak month.

C_v : Hourly weighted airport capacity.

The weighted airport capacity is computed as:

$$C_v = \frac{P_1 \times C_1 \times W_1 + P_2 \times C_2 \times W_2 + \dots + P_n \times C_n \times W_n}{P_1 \times W_1 + P_2 \times W_2 + \dots + P_n \times W_n} \quad (279)$$

Where:

$P_1, P_2, \dots, P_n$: The probabilities of use for n runway configurations.

$C_1, C_2, \dots, C_n$: The capacities for the n runway configuration.

$W_1, W_2, \dots, W_n$: The ASV weighting factor for the runway configuration.

The weighting factor is a function of the aircraft mix, weather condition and the capacity of the runway configuration relative to the maximum capacity of the airport. The variation of the weighting factor with the aircraft mix, weather conditions and the relative capacity is given in Table 3.1

Table 3.1 Variation of the ASV Weighting Factor with the Aircraft Mix and Meteorological conditions.

Percentage of Maximum Capacity	Weighting Factors			
	VFR	IFR		
		Mix Index (0-20)	Mix Index (21-50)	Mix Index (51-180)
91+	1	1	1	1
81-90	5	1	3	5
66-80	15	2	8	15
51-65	20	3	12	20
0-50	25	4	16	25

The FAA also recommends values for D and H as a function of the mix index. However, the FAA values are derived for airport terminals in the 1980's. However, the terminals have changed their operating strategy since the 1980's due to increased demand and better technologies (PRM radar), therefore the D and H factors were derived from the T-100 and the ASPM respectively. The V/C ratio for the top airports are illustrated in Table 3.2

Table 3.2 The Demand and Capacity Values for the top 33 airports in the NAS (2006).

Airport ID	ASV	V/C Ratio	Demand
ATL	990415	0.9541	944955
ORD	876972	1.0546	924855
DFW	927645	0.7294	676624
LAX	724045	0.8412	609067
IAH	625070	0.9236	577315
DEN	880309	0.6483	570704
PHX	595565	0.7906	470854
PHL	436633	1.06	462831
DTW	589539	0.778	458661
CLT	522803	0.8508	444801
MSP	559289	0.769	430093
EWR	451432	0.9494	428590
LAS	530703	0.7711	409225
LGA	363010	1.0617	385408
BOS	393221	0.9437	371083
JFK	336375	1.0987	369575
MIA	520931	0.646	336521
SEA	410616	0.8112	333092
MEM	626130	0.5266	329720
MCO	680766	0.4715	320981
CVG	669466	0.479	320674
SFO	375528	0.8493	318936
SLC	449677	0.6964	313155
IAD	462318	0.6656	307719
DCA	324558	0.8384	272109
STL	475879	0.5237	249218
BWI	409937	0.6004	246126
CLE	312293	0.7309	228255
FLL	289542	0.7627	220834
MDW	336101	0.6081	204383
PDX	446470	0.455	203144
PIT	542106	0.3711	201176
TPA	374287	0.5305	198559

The D factor is derived from the ASPM by taking the ratio of the average number of daily operations to the number of operations during the peak hour. The H factor is derived from the ratio of the number of annual operation to the average daily operations during the peak month. The ASV values derived above are for the year 2006, and therefore may not reflect changes to airport infrastructure after the year 2006. It is also necessary to consider the penalty for the utility of a route when the volume of the intermediate airport approaches capacity. Figure 3.5 shows the forms of the penalty that were considered.

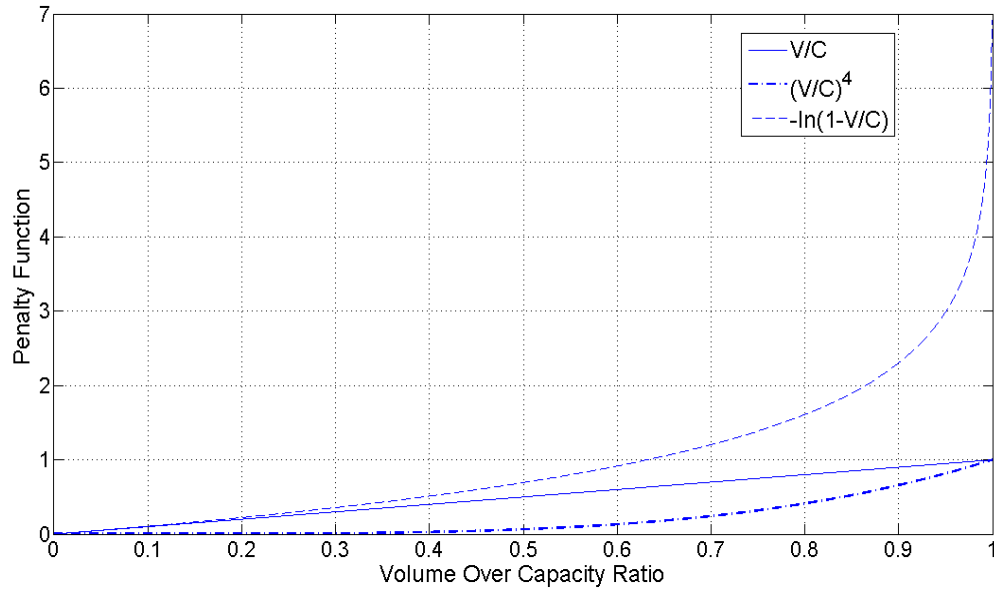


Figure 3.5 Variation of the Penalty Function with the Volume over Capacity Ratio.

Each form of the penalty function has some advantages and drawbacks. The logarithmic function imposes an absolute restriction on demand as it approaches infinity as the volume reaches capacity. However, the disadvantage is that the logarithm is undefined when the volume exceeds capacity and therefore, cannot be used in cases when the airport is operating at a demand that exceeds the capacity. The power function provides less resistance than the linear function when the demand is less than the capacity, but rapidly increases once the demand exceeds capacity thus rendering the airport very unattractive for connections when the airport is operated under over-saturated conditions. The power function is defined when the ratio of volume and capacity exceeds one and therefore avoids the shortcoming of the logarithmic function. The power function is therefore, the most suitable function for depressing demand at an airport when the demand approaches or exceeds capacity. It is also possible to use a dual-regime penalty function where the logarithmic penalty function is used when the demand is less than the capacity and the power function is used when the demand exceeds capacity. The entire calibration process is illustrated in Figure 3.6

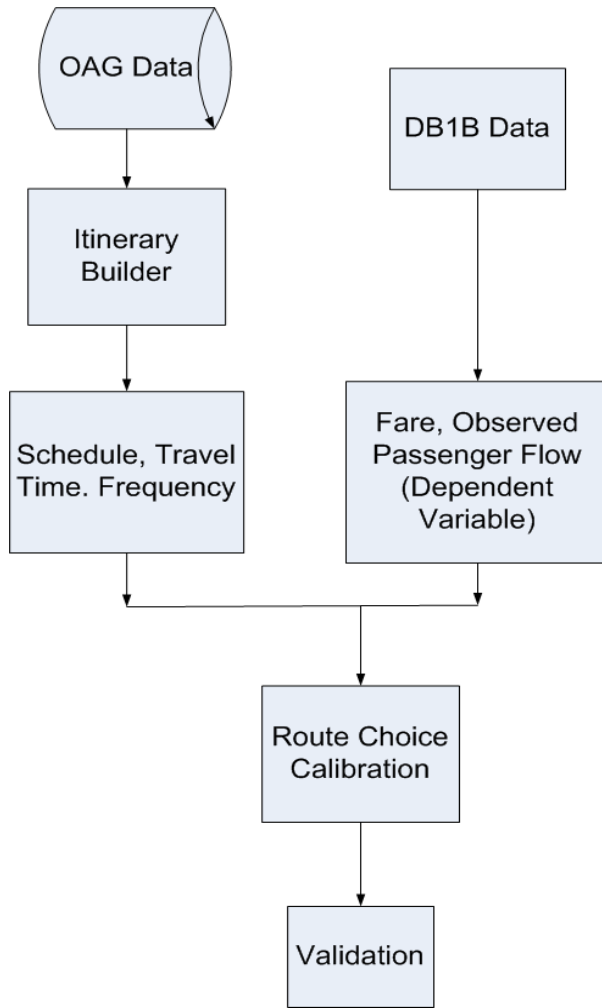


Figure 3.6 The Air Traffic Flow Problem.

3.3.2)The Schedule Generator

The Schedule derives the number of yearly flights for each aircraft type for each flight segment in the NAS. The objective of the schedule generator is to derive the cost-minimizing schedule. The constraints considered are seat, airport landside and airport airside capacities, aircraft utilization, flow conservation and runway length constraints. Mathematically, the schedule generator can be written as:

Min

$$\sum_k \sum_l N_{lk} CD_{lk} + \sum_k \sum_l N_{lk} CI_{lk} + \sum_i CL_i \sum_{l:l \in (i,j)} N_{lk} \quad (280)$$

subject to:

$$(1) \quad \text{Leg demand} \leq \text{Leg Supply} \\ \sum_m \delta_{lr} D_{rm} \leq \sum_k N_{lk} S_k LF_{lk}, \forall l \quad (281)$$

$$(2) \quad \text{Airport Demand} \leq \text{Airport Airside Capacity} \\ \sum_{j:l \in (i,j)} \sum_k N_{lk} \leq C_i = f \cdot \sum_k N_{lk} S_k, \forall i \quad (282)$$

$$(3) \quad \text{Aircraft utilization} \\ \sum_l N_{lk} * LT_{lk} \leq FT_k * OP_k, \forall a, k \quad (283)$$

$$(4) \quad \text{Flow conservation at each airport} \\ \sum_k \sum_{l:l \in (i,j)} N_{lk} = \sum_k \sum_{l:l \in (j,p)} N_{lk} \quad (284)$$

$$(5) \quad \text{Airport runway length} \\ N_{lk} = 0 \text{ for legs connecting small airports and } k=2,3. \quad (285)$$

$$(6) \quad \text{Airport Demand} \leq \text{Airport Landside Capacity} \\ \sum_{j:l \in (i,j)} \sum_k N_{lk} * S_k \leq C_i^*, \forall i \quad (286)$$

Where:

N_{lk} : Frequency of service on link l , by aircraft type k .

$k \in K$: Aircraft type $\{1, 2, 3 = \text{small, medium, large}\}$

F_{rm} : Fare charged on route r in market m .

T_{rm} : Average travel time on route r in market m .

SD_{rm} : Average schedule delay on route r in market m .

CD_{lk} : Direct operating costs of service on leg l by aircraft type k .

CI_{lk} : Indirect operating costs of service on leg l by aircraft type k .

LF_{lk} : Load factor on leg l in aircraft type k .

OP_k : Maximum operable aircraft hours allowable on fleet type k .

FT_k : Number of aircraft in fleet type k in the NAS.

LF_{lk} : Maximum Load factor on leg l for aircraft type k .

LT_{lk} : Block time on leg l for aircraft type k .

CL_i : Landing fees charged per landing by airport i .

C_i : Capacity of airport i .

S_k : Seat capacity for aircraft type k .

$\delta_{lr} = 1$ if leg l is a part of route r
 $= 0$ otherwise.

3.3.2.1) Timetable Generator

The timetable generator derives the timetable (departure and arriving times) for each flight from the output of the schedule generator. This is achieved by interpolating the CDF of departure times obtained from the OAG by the number of flights on each leg. For example, if the flight frequency from ROA to ATL is three flights per day and the flight frequency from ROA to IAD is four flights per day then the output of the timetable generator is given in Table 3.3

Table 3.3 Sample Timetable derived from the Timetable Generator.

Airline	Origin	Destination	Deptime	Arftime	Seats	Stops	Freq
XX	ROA	ATL	10:34	12:00	50	0	365
XX	ROA	ATL	15:47	17:10	50	0	365
XX	ROA	ATL	22:29	23:51	50	0	365
XX	ROA	IAD	09:12	10:12	50	0	365
XX	ROA	IAD	13:10	14:10	50	0	365
XX	ROA	IAD	17:02	18:02	50	0	365
XX	ROA	IAD	22:29	23:29	50	0	365

3.3.3) The Fare Prediction Module

The fare prediction module derives the fares for each route in the network, since the fare is an important exogenous variable to predict route flows. The fare prediction problem can be subdivided into two distinct sub-problems; the first submodule derives fares for the routes for which the DB1B has insufficient information and the second sub-module projects the baseline year fares into the future

Rama-Murthy (2003) developed a series of regression models to calculate fares based on some independent variables (distance, competition, and low-cost carrier

presence and airport type) to solve the first sub-problem. The author proposes two classes of models

- (1) Pure distance based non-linear regression
- (2) Multiple linear regression based models, considering more independent variables in addition to distance.

The fare sub-model would use the pure distance based model, since the multiple linear regression model requires information on individual airline strategy

The distance-based model is a Harris model, which can be expressed as,

$$fare_{ijk} = \frac{1}{(a + b * D_{ijk}^c)} \quad (287)$$

Where

a, b, c : Constants.

D_{ijk} : Flown distance between airports i and j through airport k .

$fare_{ijk}$: Fare between airports i and j through airport k .

Separate regression models have been developed for each airport pair type (Large, Medium, Small and Non-hub) since the service from large and medium hubs consists of a higher proportion of larger aircraft, which have lesser cost per seat-mile than smaller aircraft.

The fares are projected into the future by deriving a trend-line for fares with respect to certain explanatory variables for each distance bracket. The variables considered were:

1. Demand
2. Load Factor
3. Dummy for 9/11 terror attacks
4. Cost per seat
5. Distance

Distance was implicitly included by performing a separate trend-line analysis for each distance bracket.

3.3.4) Network Evolution Submodule

The network evolution submodule attempts to predict the evolution of the airline network topology into the future. This is necessary since airlines have changed their network structure significantly since deregulation. Airlines cut routes, which are not profitable, or do not have enough demand to justify service. On the other hand, an airline introduces direct service in markets with enough demand or to bypass congested hub airports. The network evolution submodule would also add new airports when all the available airports

in a region become saturated. The network evolution submodule can therefore be divided into three subproblems

- (i) Cutting service on non-profitable routes or routes with poor demand
- (ii) Adding point-to-point service in markets with enough demand
- (iii) Introduction of commercial service in an airport which has very little commercial service in the baseline year

3.3.4.1) Cutting service on non-profitable routes

Following the Sept 11 attacks, the airline industry posted record losses and most of the large carriers went into bankruptcy. In order to achieve profitability, the network carriers ceased services in many markets. For example, US Airways officially de-hubbed Pittsburgh International Airport in September 2004. Consequently, the number of enplanements at Pittsburgh fell by about 50% from 2003 to 2006. Therefore, a model that predicts airline business strategy is required to predict the airport demand with a good degree of accuracy. The model attempts to derive the relationship between the airline decision and certain independent variables. The independent variables used in the model are:

- (1) Profit/Seat – As was explained in the last paragraph airlines cut service in routes that have poor revenue or profit potential. Consequently the profit per seat earned by airline is one of important variables that affect an airline's strategy
- (2) Load Factor – An airline attempts to fill as many seats as possible in each flight. Any flight that is mostly devoid of passengers has no revenue potential and it is highly likely that the airline would cease to operate the flight.
- (3) Distance – Automobiles and transit compete for market share with airlines on short-haul routes. Therefore, the revenue potential for a short-haul route is less than that of medium and long haul routes and is much more unattractive for revenue generation. An airline facing losses would cease service in short-haul routes before it culls medium or long haul routes.
- (4) Frequency – The probability of service being ceased on a route is dependent on the density of service on the route. The chances of an airline ceasing service on a high-density route between two major cities are very low. On the other hand, low-yield routes, which primarily serve small communities, are routes, which are cut by the airlines, when facing financial difficulties.
- (5) Essential Air Service airports – The Essential Air Service United States Government program that was enacted to guarantee that small communities in the United States that prior to deregulation were served by certificated airlines maintained commercial service(Wikipedia(2006)). It was designed to maintain a minimal level of scheduled air service to these communities which otherwise would not be profitable. The United States Department of Transportation currently subsidizes airlines to serve approximately 140 rural communities across the country that

otherwise would not receive any scheduled air service. Therefore, an air carrier would continue service on routes, which are otherwise undesirable from an economic, and operation point-of-view and can be described as “outliers” as these routes are subsidized by the federal government. The Essential Air Service indicator variable is introduced to remove the bias in the model, which arises due to the existence of these subsidized routes.

3.3.4.1.1) Profit Computation Submodule

The profit for each route is one of the most important variables that affect the airline strategy. The profit was computed by subtracting the revenue and the cost. The cost of the routes was obtained from the operating cost of the aircraft flown on the route. The operating costs were derived from the Form 41 data.

The revenue was computed from the DB1B. However, since the DB1B is a 10% sample, the revenue for small markets could not be obtained. The market revenue obtained from the DB1B had to be divided among the flight legs in the ratio of the cost-per seat mile for each flight leg. The profit for each flight leg is computed by subtracting the cost from the revenue. For example, consider the small sample network in Figure 3.12

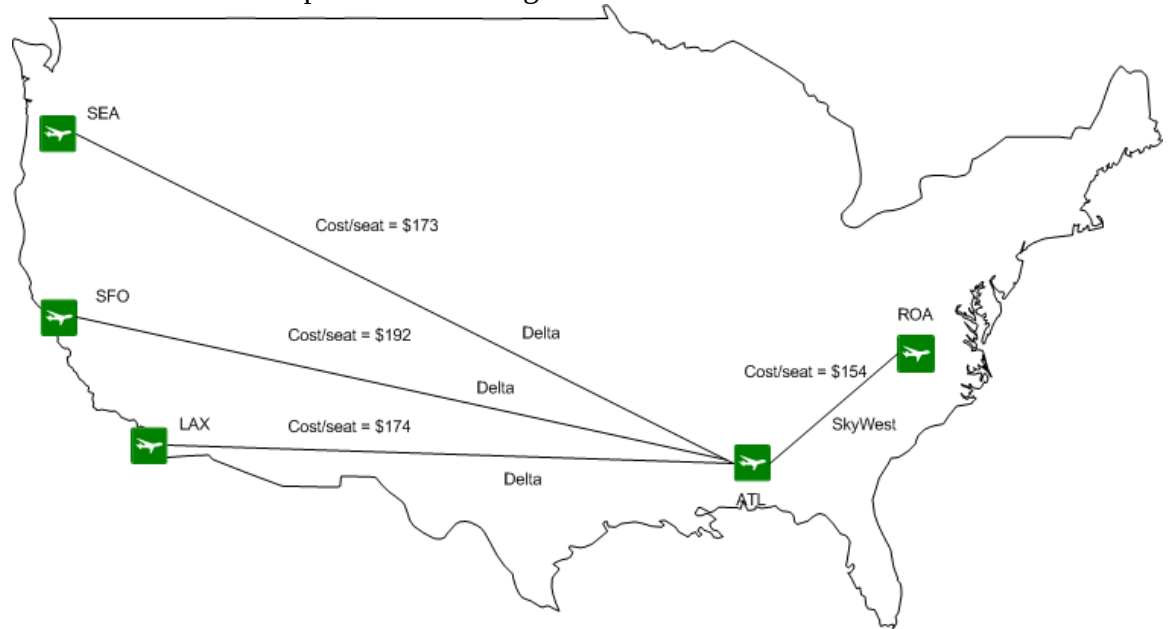


Figure 3.7 Sample Airline Network used for Revenue Computation.

The numbers in the brackets indicate the ratio of the Operating Cost per hour per seat. For example, the ratio of the Operating Cost per hour per seat of the ROA→ATL and ATL→SFO legs are in the ratio 2:3. The above ratio fails to take into account the flight times, so the ratios are modified by obtaining the ratios of the Operating Cost per seat instead of the Operating Cost per hour per seat. Assuming that the ticket revenue (DB1Bx10) for the ROA→SFO,

ROA→LAX and ROA→SEA markets are \$ 5,000,000 \$10,000,000 and \$15,000,000 respectively. Then the revenue for the ROA→ATL segment is computed as:

$$= \frac{1}{5} * 5000000 + \frac{1}{5.5} * 10000000 + \frac{1}{6} * 15000000 = \$5,318,200.$$

The cost is the sum of the operating costs of all the aircraft flying the ROA→ATL route. For example, if the Canadair CRJ-200 is the only aircraft flying the ROA→ATL route and the number of flight hours per year spent on the route is 2,000 then the cost of the ROA→ATL route is, \$ 2,000 (operating cost per hour)* 2,000 (flight-hours) = \$ 4,000,000.

The total profit on the ROA→ATL route is \$ 5,318,200-\$4,000,000 = \$ 1,318,200.

3.3.4.2) Adding Point-to-Point Service in Markets with Enough Demand

The hub-and-spoke structure was developed after deregulation. The passengers originating from a small city or a spoke were routed through a hub to their final destination. The advantage of this system was that travelers from different originating airports could be pooled into a single flight to their final destination thus leading to cost-benefit accrued from economies of scale. However, the hub-and-spoke system results in increased travel time and inconveniences to passengers due to transfers and lost luggage. Therefore, the airlines add direct service in high-density markets where there are sufficient economies of scale. Therefore, the model needs to identify the markets where the potential for direct service exists in the future. This is done in two stages. In the first stage, the model builds the set of the markets that have the potential for direct service. The threshold for direct service is 12,500 passengers per year. The threshold is derived from the minimum yearly frequency for direct service (250 flights per year) and the average aircraft size in the NAS (50 seats per flight). This threshold is not static and can change with the average number of seats per flight in the future years.

3.3.4.3) Introduction of Commercial Service in an Airport, which has Very Little Commercial Service in the Baseline Year

The continental United States has 443 airports, which have commercial service. However, the US also has about 580 airports with paved runways of length 6,000 feet or greater which could potentially serve as secondary airports when some of the 443 primary airports reach capacity.. The model selects the best candidate for developing commercial service in two stages. In first stage all the secondary airports which are more than 150 miles from the centroid of the Metropolitan Statistical Area (MSA) served by the primary airports are discarded. The secondary airport analysis is only important for MSA regions as all the congested large hubs are located in urban areas. In the second stage, the remaining candidates are ranked according to certain

criterion and best candidate airport is chosen for commercial service development. The criteria considered are:

- (1) Airport Capacity: The capacity of the reliever airport is the most important factor in the analysis, since the purpose of the reliever airport is to mitigate congestion at the primary airport. Therefore, a secondary airport with a higher capacity would ensure a greater reduction of demand at the primary airport
- (2) Distance from the MSA centroid: The secondary airport should be as close as possible to the MSA population centroid to attract the maximum possible demand. If the reliever airport is too far away from the MSA then it would attract very less demand, and fail to mitigate congestion at the primary airport
- (3) Length of the longest runway: Runway length is an important factor that could constrain the number and nature of operations at an airport. A reliever airport with a long runway could serve heavier aircraft and consequently handle higher demand.
- (4) Presence of control tower: A control tower is essential for realizing the entire capacity potential at an airport. This is because an airport with a tower can offer better traffic management and consequently have a greater capacity than a non-tower airport. A airport with a tower could also enhance the operational safety at an airport

The model assigns a rank for all the candidates based on the criterion described in the previous paragraph and the candidate with the best rank is chosen as the reliever airport. For example, the first, second and third choice for the Chicago MSA which are DuPage Airport, Gary/Chicago International Airport and Aurora Municipal Airport respectively.

Chapter 4 Model Description

4.1) Introduction

The model consists of four sub-modules

- (1) The Air Traffic Network Flow Problem
- (2) The Schedule Generator
- (3) The Network Evolution Sub-Module
- (4) The Fare Evolution Sub-Module

These four sub-modules are integrated with the Transportation Systems Analysis Model (TSAM) to obtain the trends in demand, fleet size and airport delays over the next 25 years. The sub-modules run in sequential order, in other words the output from the single sub-module becomes the output for the next model. The integration module is explained in detail in Section 4.1.1

4.1.1) Model Integration

The four sub-modules need to act in concert with the demand model in TSAM to derive meaningful trends in the state variables of the air transportation system. The principal assumption in the model is that the National Airspace System (NAS) never attains equilibrium, in a single year. The equilibrium assumption has been made by other studies that attempt to capture supply demand interaction, for example Hsiao and Hansen (2004) and Ghobrial and Kanafani (1995). However, there is a certain lag between supply and demand variables as passengers do not react to changes in supply instantaneously. Air travelers often react to changes in service quality in the next time epoch of the airline schedule planning process. For example, airlines often plan their schedule for each quarter of the year, several months in advance and typically do not favor substantial changes in schedule in the middle of the planning process. The air travelers react to the change in service and the airline changes its service based on the elasticity of the passenger demand to the change in service in the previous time period. Therefore, the model utilizes the service quality of the previous time-period as an input to determine demand for the present time period. However, for the initial and the first time periods ($t = 0$, and $t = 1$) the schedule for the initial time period ($t = 0$) is used to start the process. The entire integration process can be represented as a flowchart in Figure 4.1 and as a pseudo-code in Section 4.1.1.1

4.1.1.1) Model Integration Process

- 1) The network evolver is executed at time $t = t$ to derive the changes in the airline network topology and the initial schedule for the new segments
- 2) The schedule at time $t = t$ is updated to reflect the changes in network topology
- 3) The airline schedule at time $t = t$, is used by TSAM to obtain the airline demand (Origin-Destination Demand) for the time $t = t + 1$
- 4) The Air Traffic Network Flow Problem utilizes the airline demand derived from TSAM to derive the segment flows for time $t = t + 1$

- 5) The output of the Air Traffic Network Flow Problem is used to derive the airline schedule for the time $t = t + 1$
- 6) The network evolver is executed for time $t = t + 1$ and the process is repeated as shown in Figure 4.1

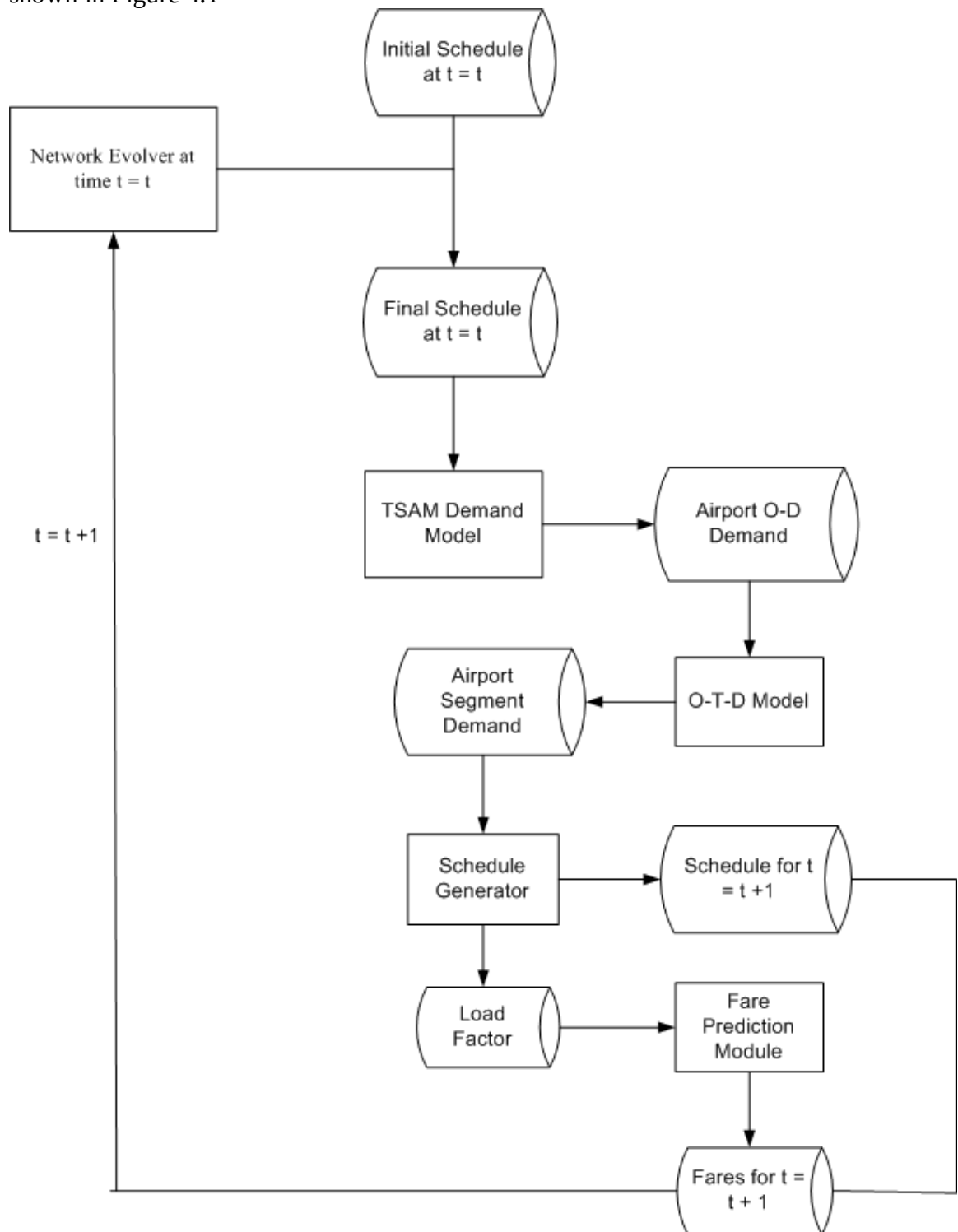


Figure 4.1 The Model Integration Process with TSAM.

4.2) The Air Traffic Network Flow Problem (O-T-D Model)

As mentioned in Section 3.3.1 the Air Traffic Network Flow Problem can be stated as,

Given a set of Origin-Destination pairs (airports), the demand between them, the set of paths covering each O-D pair and their service attributes (travel time, fares, frequency, seats) and the set of flight legs in the network. Find the demand for each flight leg, subject to link and node capacity and other additional constraints

Different methodologies were attempted for solving the problem as there have been many studies attempted on the problem of network assignment in ground transportation. However, most of the studies in ground transportation utilize the concept of user equilibrium, which is not applicable for airline networks since the service providers in ground networks (roadways, vehicles) are passive agents unlike the service providers (airlines) in an airline network. However, some studies have been attempted to solve the network assignment problem in airline networks. The main methodologies used are logit models, game theoretic models, fuzzy set based multi-agent models and network equilibrium models.

The logit model is based on the premise that every decision-maker maximizes his utility when presented with a range of choices. The utility is usually expressed as a linear function of several independent variables. The coefficients for the independent variables are estimated using a maximum likelihood technique. The logit model is computationally the simplest of all the three methodologies. Logit based formulations are employed by Weidner (1995), Ghobrial and Kanafani (1995) and Hsiao and Hansen (2005). Weidner (1995) uses a nested logit model to solve the problem of passenger route choice. Supply costs, congestion, and other airport characteristics are used to model the route utility. Route distance and capacity saturation are used as surrogates for travel time and connecting delays. Ghobrial and Kanafani (1995) use a logit based equilibrium model to determine segment flows for 25 airports in the Southeastern region of the United States. The independent variables considered are travel time, fare, frequency, number of legs and seats. The model initially assumes unit frequency for determining segment demand. A pre-set load factor was then assumed for computing frequency on each leg. The process is repeated until equilibrium was achieved. Hsiao and Hansen (2005) use a logit-based equilibrium formulation to model segment demand. The variables used in the model are similar to Weidner (1995). The authors use delays instead of capacity. Assuming an initial delay, the process is iterated until equilibrium is reached. However, the logit model cannot take into account the entire spectrum of factors in the segment flow problem, for example, airline competition whose dynamics cannot be captured in the logit model and airline bankruptcies, which are difficult to model.

Game Theory models competition between airlines and passengers are passive agents in the game. Game Theory can address the issue of competition between airlines. Game-theoretic formulations are used by Hansen (1990), Dobson and Lederer (1993) and Adler (2005). Hansen (1990) uses an n-player non-cooperative game to determine airline shares and segment flows. The variables used to model route utility are fare, route circuitry and

the maximum and minimum flight frequency on each link of the route. Dobson and Lederer (1993) use a game-theoretic framework to schedule flights and calculate segment flows. The route demand is determined by assigning a disutility to each itinerary for that route. The authors then maximize profit for each airline, which is a function of the airline schedule. The profit-maximizing schedule is obtained and therefore, the profit and segment flows. Adler (2005) uses a game-theoretic framework to model airline competition in Western Europe. The model scope is 20 nodes and four airlines. However, implementing game theory on a large network is quite difficult. Dobson and Lederer (1993) state that a parallel computer with several thousand processors is required for implementing game theory on a large (50-node) network. All existing literature on the application of game theory only deal with networks of 10-30 cities which is about an order of magnitude less than the problem dealt with in this analysis.

Agent-based formulations are used by Teodorovic and Kalic (1994) and Neidringhaus (2004). The authors use a set of rules based on fuzzy set to incorporate passenger choice in airline networks. However, precise formulation of the rules for a large-network can be quite involved because of the number of interacting phenomena. Teodorovic and Kalic implemented their model on a network consisting of 13 nodes in the Southeastern part of the United States. Niedringhaus implements the jet-wise agent based simulation model on all the Metropolitan Statistical Areas (MSAs) in the Continental United States.

Network equilibrium models assume that no passenger can change his route without decreasing his utility. Soumis and Nagurney (1993) propose a multi-class stochastic network equilibrium model to predict passenger route choice. The authors solve the network equilibrium problem as a series of simultaneous non-linear equations and obtain the final route flow in the market. The model is applied to the Air Canada Network. However, air transportation networks may not obey equilibrium conditions due to the more complicated interactions between the service providers and the passengers. The methodology finally adopted is the logit model since it is not clear that the other methodologies could be implemented for a large-scale problem and the logit model offered computational simplicity, which was the quickest method to achieve an acceptable solution from the four methods.

The model consists of three sub-modules: 1) A schedule-building module, which derives a representative yearly schedule; 2) A network-building module, which constructs a synthetic network using the OAG data; and a 3) Choice module, which derives the coefficient set for the explanatory variables. The schedule generator and the network generator module have been explained in Section 3.3.1.1. In this section, all the three modules will be explored in greater detail

4.2.1) The Schedule Building Module

The schedule building module generates a representative yearly schedule from the Official Airline Guide (OAG) data. The schedule building module smoothes out seasonal variations in the airline schedule. The model does not take into account seasonal

variations since it is a yearly planning model. Therefore, eliminating seasonal variations would ensure that removal of any bias in the input variables. The schedule builder constructs an effective yearly schedule by assuming that flights by the same airline between the same O-D pair, which arrived or departed within 30 minutes of each other in the OAG are the same flight. The time-period of analysis for the schedule builder is October 31st to March 2nd due to the daylight savings time. Any flight with an effective frequency of less than 5days/week is eliminated. The schedule builder module is depicted as a flowchart in Figure 4.3 and a sample of the result from the module is shown in Figure 4.2.

	carrier1	flightno1	depairport	arrairport	lcdeptime	localartm	seats	efffrom	effto	elapsedtme	flyngtime	groundtime	stops	frequency
1088	AA	532 ABQ	ORD		835	1216	133	6/15/2000	7/1/2000	241	241	0	0	17
1089	AA	532 ABQ	ORD		835	1216	133	8/1/2000	8/5/2000	241	241	0	0	5
1090	AA	532 ABQ	ORD		835	1216	133	8/6/2000	9/5/2000	241	241	0	0	31
1091	AA	1104 ABQ	ORD		1125	1510	133	7/2/2000	7/31/2000	245	245	0	0	30
1092	AA	1104 ABQ	ORD		1130	1515	133	7/1/2000	7/1/2000	245	245	0	0	1
1093	AA	1104 ABQ	ORD		1130	1515	133	8/6/2000	9/5/2000	245	245	0	0	31
1094	AA	1104 ABQ	ORD		1130	1515	133	6/15/2000	6/30/2000	245	245	0	0	16
1095	AA	1104 ABQ	ORD		1130	1515	133	8/1/2000	8/5/2000	245	245	0	0	5

	depairport	arrairport	carrier1	stops	flightno1	frequency	lcdeptime	localartme	elapsedtme	flyngtime	groundtime	seats
214	ABQ	ORD	AA	0	532	365	830	1210	240	240	0	133
215	ABQ	ORD	AA	0	1104	365	1141	1524	243	243	0	133

Figure 4.2 The Schedule Building Module.

It should be noted that the schedule building module is only necessary for the baseline year, where the schedule is taken from the Official Airline Guide (OAG). The schedule for the future years would be derived from the Schedule Generator and therefore, the schedule building module would not be required.

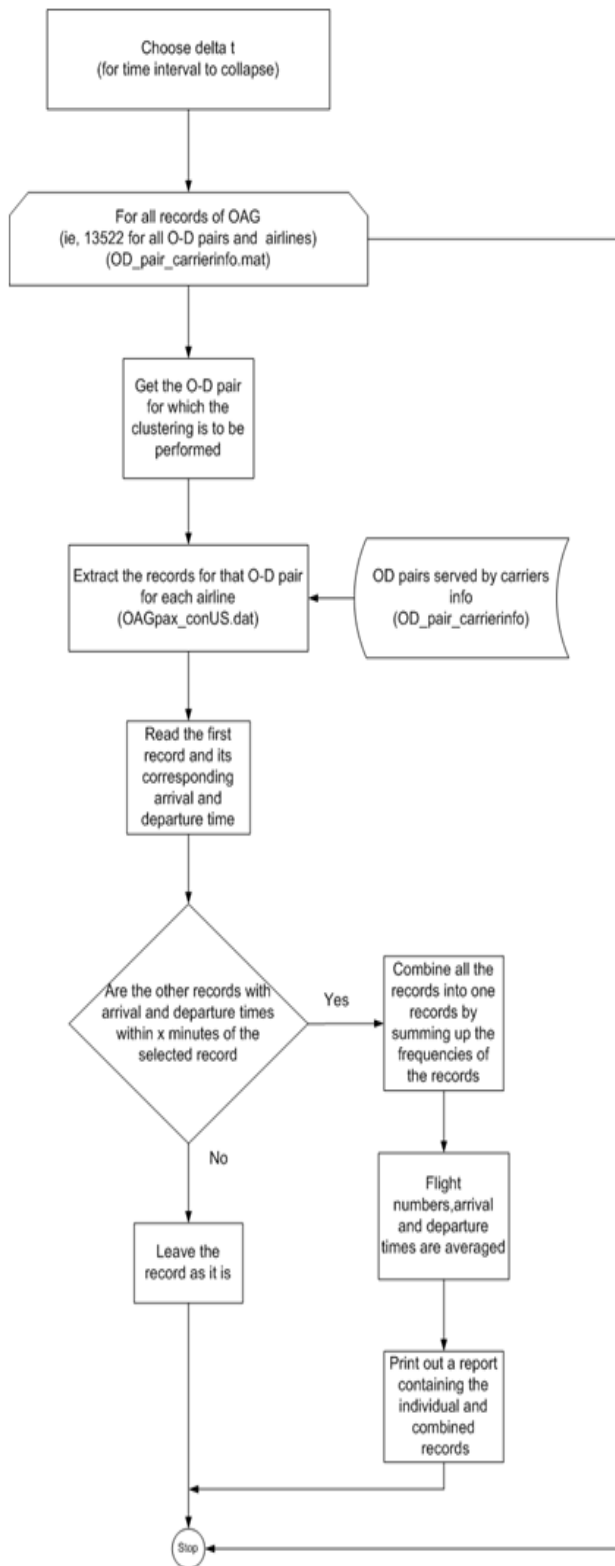


Figure 4.3 The Schedule Building Module.

4.2.2) The Network Building Module

The network building module has been already explained in detail in Section 3.3.1.1. However, the network building heuristic leads to multiple paths between O-D pairs. In section, the heuristic to discriminate between alternate paths between the same O-D pair are explained.

4.2.2.1) Rules for the Network Building Module

Rule 1 (Same airline): Flying on the same airline is preferred to switching carriers, since an airline would desire to preserve its market share. Switching airlines is considered when the difference in travel time was considerable (minimum of six hours or 50% of the larger travel time).

Rule 2 (1 leg vs. 2-leg): In the previous iteration it is assumed that if there are direct flights between two airports, all or most passengers would take only direct flights. However this assumption was not always true. For example analysis of ROA-ORD segment illustrated that many of the passengers connected through CLT, ATL or some other hub. Therefore, all possible 2-leg itineraries on the same airline are appended to the direct flights

Rule 3 (2 leg vs. 3-leg): Initially it is assumed that two-leg flights are always better than three-leg ones. This is not always true, for example a trip from HOT (Memorial Field, TX) to BIL (Billings Logan Intl, MT), a two-leg solution gives a travel time of 15 hours through DFW. A three-leg solution through DFW and DEN gives a travel time of 8 hours. Therefore a two leg solution is replaced by a three-leg solution whenever the difference in travel times was considerable (minimum of six hours or 50% of the longer travel time).

Rule 4 (3-leg vs. 4 and more legs): The 3-leg solution is replaced by the solution with 4 or more legs if the sum of the travel time and schedule delay for the 4 or more legs solution is less than the corresponding sum of travel time and schedule delay for the 3-leg solution

Rule 5 (backward skim for scheduling): For flights greater than 3 legs, it is computationally intractable to derive the optimal combination of flight legs for the minimum travel time. Therefore the connecting time was minimized in stages, with the assumption that the solution that is optimal in stages would yield the global optimum. It is found that this assumption gave a very bad solution when the one of the intermediate leg flights was very late in the day, or one of the legs had a very low frequency, leading to a large connecting time of the order of 24 hours. For these O-D pairs, the scheduling is started backwards, with the final leg being flown late in day so that sufficient time was allowed for connections on the other two legs, resulting in a considerably lower travel time.

4.2.3) The Discrete Choice Module

The Choice module determines the probability of choosing a particular route between an origin-destination airport pair. As explained in Section 2, there are various methods that can accomplish this. The logit model is selected to model passenger route choice based on computational simplicity and explanatory power. The variables used in the model are described in Table 4.1.

Table 4.1 Variables Used in the Logit Formulation.

Variable	Description	Reason for introduction
Relative Travel Time	Quotient of the Route Travel Time and the Average O-D pair Travel Time	Travel Time is an important factor in determining the utility of a route
Relative Fare	Quotient of the Route Fare and the Average O-D pair Fare	Travelers are sensitive to fare charged on a flight
Seats	The number of seats on the first leg on a route	The number of seats offered is a measure of the probability of a passenger being denied a seat on the flight
Legs	The number of legs on the route	Passenger avoid connecting flights whenever possible
Dummy Variable	Airport specific dummy variables	To capture airport specific phenomenon not accounted for in the other variables
Delays	Average delay per operation in an airport	Passengers avoid airports where they experience significant delays

Owing to the various geographical, demographic and various traveler characteristics, it is necessary to derive a scheme to segment the input data. The various segmentations that are considered are described in Table 4.2.

The data was segmented by the number of legs. Therefore, two logit models are calibrated 1) one for O-D pairs linked by both direct and connecting itineraries and 2) the second logit was for O-D pairs linked by connecting itineraries alone. Three separate logit models were calibrated:

- (i) Multinomial logit without the airport delay variable.
- (ii) Multinomial logit with the airport delay variable.
- (iii) Nested logit with the airport delay variable for markets served by both direct and connecting itineraries and a multinomial logit for markets served by connecting itineraries alone.

Table 4.2 Segmentations Considered in the Logit Formulations.

Segmentation type	Description	Reasons for consideration	Reasons for inclusion/non-inclusion
Hub Type	The hub type of the origin and destination airport according to the FAA definition	To account for difference in passenger behavior across airport type	Did not significantly enhance the predictive power of the model
Ticket Type	The ticket class (business class or economy class)	To account for the fare structure variation across ticket classes.	Fare normalization accounts for fare structure variation
Number of Legs	The number of legs in the itinerary.	To account for passenger choice difference when a direct flight is offered	Wiedner (1995) and Hsiao and Hansen (2005) calibrate two separate models for O-D pairs with and without direct flights
Trip purpose	The purpose of the trip	Differing sensitivity for Travel Times and Fares for business and leisure travelers	Could not obtain trip purpose from the DB1B.

The utility functions used in the model can be expressed as:

$$U_{ijk} = \alpha_1 * RTT_{ijk} + \alpha_2 * RTC_{ijk} + \alpha_3 * Freq_{ijk} + \alpha_4 * S_{ijk} + \alpha_5 * \eta_{ijk} + \beta i \quad (288)$$

$$U_{ijk} = \alpha_1 * (RTT_{ijk} + RDel_k) + \alpha_2 * RTC_{ijk} + \alpha_3 * Freq_{ijk} + \alpha_4 * S_{ijk} + \alpha_5 * \eta_{ijk} + \beta i \quad (289)$$

Where:

i : The origin airport.

j : The destination airport.

k : The intermediate airport.

RTT_{ijk} : The relative route travel time.

RTC_{ijk} : The relative route fare.

$Freq_{ijk}$: The number of feasible itineraries on the route.

S_{ijk} : The number of seats offered by the airline from airport i to airport k .

η_{ijk} : The number of legs on the itinerary.

$RDel_k$: The relative delay at airport k .

β_i : Dummy variable for some airports i .

The route passenger flow can be expressed for the multinomial logit model as:

$$D_{ijk} = D_{ij} * \frac{\exp(U_{ijk})}{\sum_k \exp(U_{ijk})} \quad (290)$$

Where:

D_{ij} : The passenger demand from airport i to airport j .

D_{ijk} : The passenger demand from airport i to airport j through airport k .

Three versions of the logit model were calibrated , two were multinomial logit models with and without the inclusion of airport delay and the third model was the nested logit model with the inclusion of airport delay. The best model is selected based on the optimal balance of computational complexity and predictive power. The best model is the multinomial logit with the airport delay variable as the model was sensitive to the capacity constraint at the connecting airport. The nested logit model did not have a significantly better log-likelihood ratio and the inclusive values for the nested logit are close to unity. Therefore, it is concluded that a nested logit would not yield any appreciable increase in the predictive power of the model. A mixed logit calibration was attempted, however it was found that the mixed logit calibration procedure was computationally intractable for even moderately large (500,000 passengers) datasets.

Several issues had to be addressed in the model. The first issue is the lack of adequate fare observations for certain routes in the network. The DB1B is only a 10% sample of all the tickets sold by the airlines in the United States. The DB1B is therefore a weak data sample for smaller airports and thinner routes. However, since fare is a critical explanatory variable for the utility function in the logit model, a method to derive a credible representative route fare is needed. This is accomplished using a non-linear regression in which the dependent variable is the fare-per-mile and the independent variable is the route distance. The equation is used in-lieu of the actual fare when adequate fare data did not exist for the route. The second issue is the overestimation and underestimation of enplaning passengers at certain large-hubs by more than 10%. Specifically, the model overestimated the enplanements at IAD by about 16% and underestimated the enplanements at ATL by about 11%. To overcome this problem, the model incorporates dummy variables for the airports, to minimize prediction errors.

The first step in the calibration procedure is the preparation of input data. This involved collapsing of the OAG schedule to derive a representative yearly schedule. The fare data from the DB1B also has to be cleaned, as the data has a number of inconsistencies. For example, the DB1B reports all fares from Southwest Airlines as first class fares, which is clearly erroneous as Southwest Airlines has no class distinction. The next step involves preparation of the calibration data set. As explained in the previous paragraph, two logit models are calibrated, one for airport pairs with connecting flights only and another model for airport pairs with direct and connecting flights. Therefore, two sets of input data are selected randomly from the DB1B. The travel time, seats and frequency are

obtained from the network building module described in Section 4.2.2. The nested logit calibration procedure is executed in SAS, and the logit model coefficients are obtained.

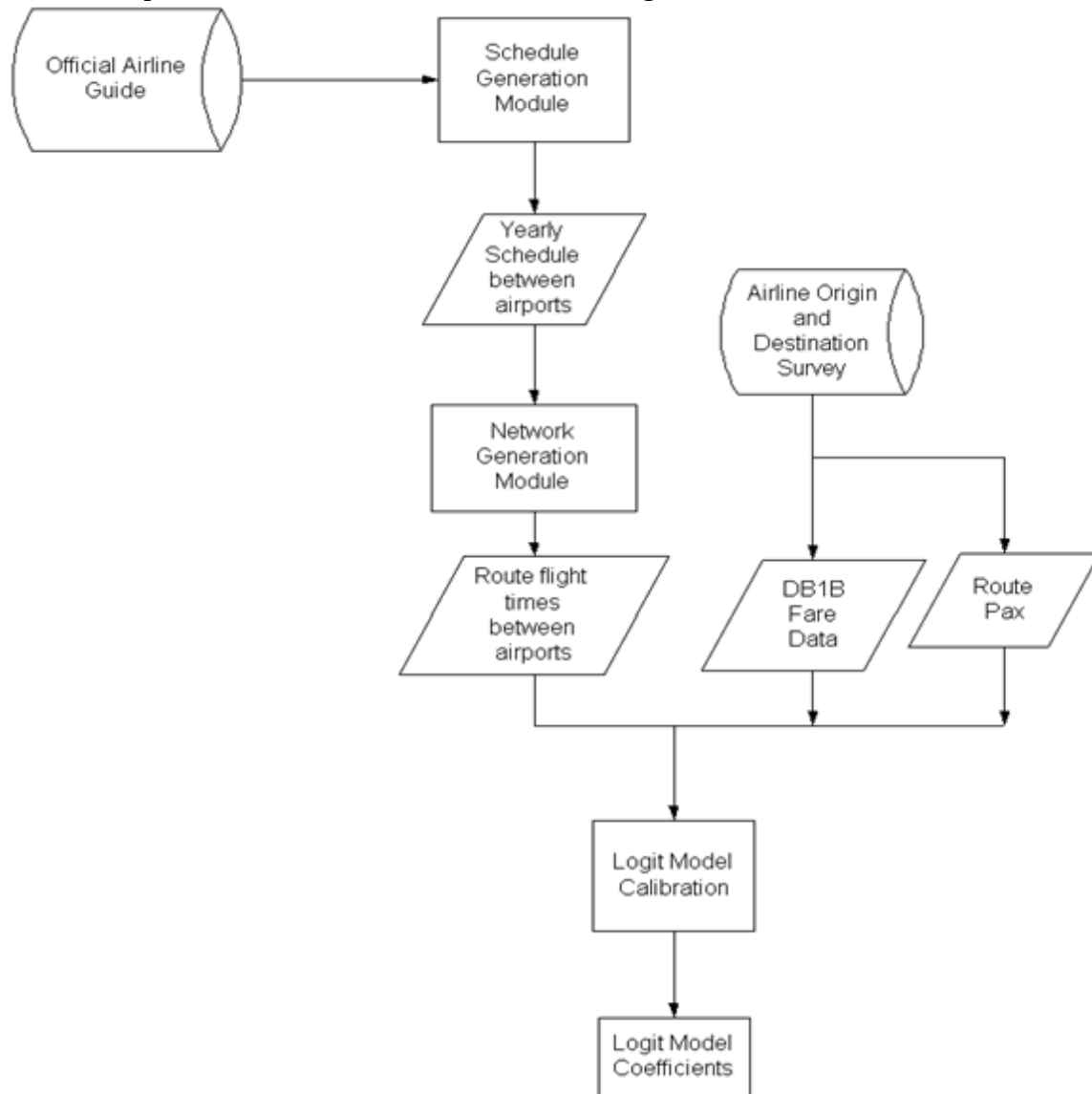


Figure 4.4 Calibration Procedure in the O-T-D Analysis.

The airport delays were derived from the analysis described in Section 3.3.1.3. The calibration procedure is illustrated in Figure 4.4

4.3) Schedule Generator

The scheduling algorithm described in this paper is used to construct a yearly schedule, given the segment demand. Two approaches to the schedule building process are possible, a clean slate approach that constructs a fresh optimal schedule, disregarding the schedule for the previous year and an incremental procedure that modifies the schedule of the previous epoch. The second approach is more suitable for airlines, as airlines cannot

accommodate large changes in frequencies and aircraft types between two time periods. Nevertheless, both approaches are attempted and their accuracy with respect to the real schedule is compared.

4.3.1) Aircraft Type Classification

The T-100 Segment has about 80 aircraft types in the Continental United States. It is clear that modeling each of these aircraft types would be superfluous since many of them share very similar characteristics. For example, the operating cost and the seats per single aircraft for a Boeing 737-300 and the Airbus A-320-200 are very close and modeling these two aircraft individually will lead to increase the size of the problem and introduce computational inefficiency. Therefore, the 80 aircraft types in the T-100 segment are classified into clusters based on the number of seats and the hourly operating costs. The operating cost for each aircraft is obtained from the DOT Form 41 database. The aircraft clusters along with their average operating cost and seats per aircraft are given in Table 4.3. From Table 4.3 it can be seen that the same aircraft type can span multiple clusters. For example, the Boeing-737 spans three classes depending on the particular generation of the aircraft. The number of seats and the operating cost for each class is the weighted average of the number of seats and operating cost for each aircraft class. Any new aircraft introduced in the future, for example the Airbus A-350 or Boeing 787 could either be accommodated in the clusters described above or a new cluster for the newer generation aircraft could be created.

4.3.2) Through Flights

The Airline Schedule in 2005 has a significant proportion of through flights. A through flight is defined as a flight where the same aircraft makes multiple stops, and passengers could enplane and deplane at each stop along the route. Southwest Airlines has a high percentage of through flights compared to other airlines as shown in Table 4.4. The model needed to account for these flights as existing literature in airline scheduling do not take through flights into account. The model uses a postprocessor to calculate the frequency of the through flights. The flights are then scheduled by concatenating the schedule of the individual flight segments.

Table 4.3 Hourly Operating Cost and Seats Per Aircraft for Each Aircraft Cluster.

Cluster Number	Operating Cost Per Hour (\$ per Hour)	Average Seats per Aircraft	Sample Representative Aircraft
1	304	9	Beech King Air, Cessna 208, Cessna C-402
2	541	19	Beech 1900 A/B/C/D, BaE J-31, BaE-748
3	1596	45	Canadair Rj-100/200, Embraer 135, Embraer 145, Fokker F-28
4	2045	72	Fokker 100, Embraer 170, Embraer 190
5	3015	127	Boeing 737-100/200, Boeing 737-300, Airbus A-319
6	2921	152	Boeing 737-800/900, Airbus A-321
7	4717	119	Boeing 737-200c, Boeing 727-200
8	4300	189	Boeing 757-200/300
9	4899	219	Boeing 767-200/300/400
10	6868	272	Boeing-777/747-400, Airbus A-300/330.

Table 4.4 Percentage of Stopover Flights for the Major Airlines in 2004.

Airline	Percentage of Stopover Flights
United Airlines	13.38
US Airways	19.11
American Airlines	5.02
Delta Airlines	20.45
Continental Airlines	13.73
Northwest Airlines	15.50
Southwest Airlines	41.02
America West Airlines	25.90
JetBlue Airlines	6.26
AirTran Airways	7.09
Frontier Airlines	24.96
Alaska Airlines	11.86
ATA Airlines	9.26

4.3.3) Model Description

The objective of the model, as described in Section 4.3 is to derive the yearly schedule for the continental United States, given the segment demand and fleet size. There are two ways to solve this problem

- (1) The new schedule can be built from scratch without any regard to the operational practices of the airlines in the previous time epoch (2004)
- (2) The new schedule can be constructed by modifying the airlines operating schedule from the previous time epoch. This approach explicitly accounts for the airlines previous operational practices.

In this paper both approaches are presented and their benefits and drawbacks are discussed. The model can be divided into four sub-modules

- (1) Frequency Generator: This sub-module derives the optimal solution of the linear programming relaxation of the original problem
- (2) Integer Solution Generator: The Integer Solution Generator derives an integer solution from the output of the frequency generator
- (3) Daily Frequency Generator: The daily frequency derives an optimal equivalent daily frequency from the output of the integer solution generator
- (4) Schedule Generator: The schedule generator is the final step in the model. The schedule derives the daily operating schedule.

(5) Postprocessor to schedule through flights

4.3.3.1) Frequency Generator

As explained in Section 3, the frequency generator derives the optimal solution to the LP relaxation of the original Integer Program. The inputs to the frequency are the demand, aircraft operating cost and fleet size and the output is a real number, which is the optimal LP solution for the integer programming problem. Various models are examined and the formulation which minimizes the sum-of-squared error between the actual and computed frequency is proposed for implementation. The first three approaches attempt to construct the schedule from scratch and the fourth formulation attempts to build the schedule by adjusting the schedule of the previous year. The constraints considered in the first three formulations are:

1. **Seat Capacity constraints:** The segment demand should be satisfied by the seat capacity offered on the segment.
2. **Minimum load factor constraint:** The seat capacity offered on the segment should not exceed the segment demand times the reciprocal of the minimum allowable load factor.
3. **Airport capacity constraint:** The annual number of operations at an airport cannot exceed its annual service volume.
4. **Fleet utilization constraint:** The total number of aircraft hours used in the schedule cannot exceed the available aircraft hours for each aircraft type.
5. **Runway length constraint:** Heavier aircraft cannot take-off or land in airports that have short runways.
6. **Airline market constraints:** Segments with lower demand, in general are not served by the mainline carriers. Therefore, markets with low demand cannot be served by larger aircraft typically used by mainline carriers.
7. **Payload range constraint:** Small aircraft cannot be used to service segments with larger stage lengths.
8. **Minimum frequency constraint:** The schedule has to maintain a minimum acceptable service frequency for each segment.
9. **Maximum frequency constraint:** The frequency for a particular segment cannot exceed a maximum allowable value.
10. **Flow conservation constraint:** The number of aircraft departing and arriving at an airport should be equal.

4.3.3.1.1) List of Variables Used

Decision Variables

N_{ijk} : Frequency of service from the origin airport i to the destination airport j .

ΔN_{ijk} : Change in frequency of service from the origin airport i to the destination airport j of aircraft k .

y_k : The number of aircraft procured.

z_k : The number of aircraft salvaged.

Model Parameters

$\Delta N_{ij}^{\max}$: Maximum allowable change in frequency of service from the origin airport i to the destination airport j of aircraft type k .

D_{ij} : Annual demand from the origin airport i to the destination airport j .

ΔD_{ij} : Change in annual demand from the origin airport i to the destination airport j .

d_{ij} : The flight distance in miles between the origin airport i and the destination airport j .

S_k : Seating capacity for aircraft of type k .

$(LF_{ij})_{\max}$: The maximum load factor between origin airport i to the destination airport j .

$(LF_{ij})_{\min}$: The minimum load factor between origin airport i to the destination airport j .

T_{ij} : Flying time from origin airport i to the destination airport j .

F_k : The number of aircraft of type k .

C_k : The hourly cost of operation for aircraft of type k .

V_i : The annual service volume of airport i .

H_k : The allowable daily use (in hours) of aircraft type k .

R_i : Length of the longest runway of airport i in feet.

$N_{ij}^{\min}$: Minimum frequency for the segment from i to j .

$N_{ij}^{\max}$: Maximum frequency for the segment from i to j .

S_{ij} : The number of seats offered from the origin airport i to the destination airport j in the baseline T-100 schedule.

$\overline{N}_{ijk}$: The number of flights offered from the origin airport i to the destination airport (j) in the baseline T-100 schedule.

Y_k : The procurement cost for each aircraft.

Z_k : The salvage cost for each aircraft.

$\sum_k^{\max}$: The maximum possible number of aircraft procured.

$\sum_k^{\max}$: The maximum possible number of aircraft salvaged.

4.3.3.1.2) Formulation I

Formulation I minimizes the operating cost of the schedule, subject to demand, capacity and fleet utilization constraints. The formulation also accounts for aircraft payload-range, runway length and load factor constraints. The variables used in the model are:
Formulation I can be written mathematically as:

Minimize

$$\sum_i \sum_j \sum_k N_{ijk} T_{ij} C_k \quad (291)$$

subject to:

$$(LF)_{\min} \sum_k N_{ijk} S_k \leq D_{ij} \leq (LF)_{\max} \sum_k N_{ijk} S_k \forall i, j \quad (292)$$

$$\sum_j \sum_k N_{ijk} + \sum_j \sum_k N_{jik} \leq V_i, \forall i \quad (293)$$

$$\sum_j \sum_i N_{ijk} T_{ij} \leq F_k H_k, \forall k \quad (294)$$

$$N_{ijk} = 0, \forall i, j \text{ where } R_i, R_j < 7000 \text{ and } k \geq 5 \quad (295)$$

$$N_{ijk} = 0, \forall i, j \text{ where } D_{ij} < 33000 \text{ and } k \geq 5 \quad (296)$$

$$N_{ijk} = 0, \forall i, j \text{ where } d_{ij} > 1,000 \text{ and } k = 1, 2, 3 \quad (297)$$

$$\sum_k N_{ijk} \geq N_{ij}^{\min}, \forall i, j \quad (298)$$

$$\sum_j N_{ijk} - \sum_j N_{jik} = 0, \forall i, k \quad (299)$$

$$N_{ijk} \geq 0, \text{ and } N_{ijk} \text{ integer} \quad (300)$$

The constants in equations 295, 296 and 297 are derived from the properties of the current operating schedule in the T-100.

4.3.1.1.3) Formulation II

Formulation I attempted to minimize the total operating cost of the schedule. However, the service frequency was not adequate from the traveler's perspective. Therefore, the second formulation attempted to maximize the service frequency, since this is more desirable from the traveler's perspective. Formulation II can be written mathematically as:

Maximize

$$\sum_i \sum_j \sum_k N_{ijk} \quad (301)$$

subject to

(262) - (270)

The constants in equations 295,296 and 297 are derived from the properties of the current operating schedule in the T-100.

4.3.1.1.4) Formulation III

Formulation I under-predicted the flight frequencies and resulted in the underutilization of smaller aircraft and Formulation II resulted in excess operations in the system and the under utilization of heavier aircraft. Formulation III attempts to optimize the objectives of the passenger and the airline simultaneously. Formulation III can be written mathematically as:

Maximize

$$\sum_i \sum_j \sum_k N_{ijk} \quad (302)$$

subject to

$$\sum_i \sum_j \sum_k N_{ijk} T_{ij} C_k \leq (1 + \Delta) \pi^* \quad (303)$$

(262) – (270)

Where π^* is the minimum operating cost obtained from solving Formulation I. The constants in equations 295,296 and 297 are derived from the properties of the current operating schedule in the T-100.

4.3.1.1.5) Formulation IV

The formulations presented in Sections 4.3.1.1.1, 4.3.1.1.2 and 4.3.1.1.3 attempt a clean-slate approach to the schedule building problem, by disregarding, the schedule for the previous time period. However, the optimal schedule could vary significantly between successive time periods, necessitating a large adjustment in airline fleet and maintenance bases. In fact as evidenced by Figures 5.10-5.12 this approach is not adequately representative in practice.

Therefore, formulation IV builds the schedule for the succeeding year based on the schedule for the previous year. The formulation also limits the allowable changes in schedule to ensure that airlines do not have to implement large changes in fleet and maintenance bases. The objective function minimizes the incremental cost accrued due to the change in schedule.

The constraints considered in formulation IV are:

1. **Airport capacity:** The annual number of operations at an airport cannot exceed its annual service volume.
2. **Fleet utilization:** The total number of aircraft hours used in the schedule cannot exceed the available aircraft hours for each aircraft type.
3. **Runway length:** Heavier aircraft cannot take-off or land in airports that have short runways.
4. **Airline market niche:** Segments with lower demand, in general, are not served by the mainline carriers. Therefore, markets with low demand cannot be served by larger aircraft used by mainline carriers.
5. **Payload range constraints:** Small aircrafts cannot be used to service segments with relatively larger stage lengths.
6. **Aircraft type feasibility:** The total frequency reduction in a segment for each aircraft type cannot be greater the service frequency of the aircraft type on the segment.
7. **Maximum and minimum segment load factor:** The seat capacity offered on the segment should not exceed the segment demand times the reciprocal of the minimum allowable load factor, and should not be less than the segment demand times the reciprocal of the maximum allowable load factor.
8. **Minimum and maximum segment frequency:** The segment frequency cannot be less than a certain minimum frequency and cannot exceed a maximum allowable value.
9. **Allowable frequency variations:** The variations in segment frequency between two successive time periods cannot lie outside an allowable bound.
10. **Flow conservation:** The number of aircraft departing and arriving at an airport should be equal.

Mathematically formulation IV can be written as follows:

Minimize

$$\sum_i \sum_j \sum_k C_k T_{ij} \Delta N_{ijk} \quad (304)$$

subject to:

$$\sum_j \sum_k \Delta N_{ijk} + \sum_j \sum_k \Delta N_{jik} \leq V_i - \sum_j \sum_k N_{ijk} - \sum_j \sum_k N_{jik}, \forall i \quad (305)$$

$$\sum_i \sum_j \Delta N_{ijk} T_{ij} \leq \Delta(F_k H_k), \forall k \quad (306)$$

$$\Delta N_{ijk} \leq 0, \forall i, j \text{ where } R_i, R_j \leq 7,000 \text{ and } k \geq 5 \quad (307)$$

$$\Delta N_{ijk} \leq 0, \forall i, j \text{ where } R_i, R_j \leq 5,000 \text{ and } k = 4 \quad (308)$$

$$\Delta N_{ijk} \leq 0, \forall i, j \text{ where } D_{ij} \leq 33,000 \text{ and } k \geq 5 \quad (309)$$

$$\Delta N_{ijk} \leq 0, \forall i, j \text{ where } d_{ij} > 1,000 \text{ and } k = 3 \quad (310)$$

$$\Delta N_{ijk} \leq 0, \forall i, j \text{ where } d_{ij} > 200 \text{ and } k = 1 \quad (311)$$

$$\Delta N_{ijk} \leq 0, \forall i, j \text{ where } d_{ij} > 640 \text{ and } k = 2 \quad (312)$$

$$\Delta N_{ijk} \leq 0, \forall i, j \text{ where } d_{ij} > 1,350 \text{ and } k = 4 \quad (313)$$

$$\Delta N_{ijk} \geq -N_{ijk}, \forall i, j, k \quad (314)$$

$$\frac{D_{ij}}{(LF_{ij})_{\max}} \leq S_{ij} + \sum_k \Delta N_{ijk} S_k \leq \frac{D_{ij}}{(LF_{ij})_{\min}}, \forall i, j \quad (315)$$

$$N_{ij}^{\min} \leq \sum_k (N_{ijk} + \Delta N_{ijk}) \leq N_{ij}^{\max}, \forall i, j \quad (316)$$

$$-\Delta N_{ij}^{\max} \leq \sum_k \Delta N_{ijk} \leq \Delta N_{ij}^{\max}, \forall i, j \quad (317)$$

$$\sum_j \Delta N_{ijk} - \sum_j \Delta N_{jik} = 0 \forall i, k \quad (318)$$

$$\Delta N_{ijk} \text{ integer} \quad (319)$$

The constants in equations (307) - (313) are derived from the properties of the current operating schedule in the T-100.

4.3.1.1.6) Formulation V

Formulations I-IV assume a constant fleet size and fleet composition. However, the schedule generator has to generate the schedule for the future years, when the fleet composition and fleet size is dynamic. The airlines purchase or sell aircraft depending on their service requirements. Aircraft that are procured or salvaged have an associated procurement and salvage cost, and airlines would try to minimize the sum of these costs, assuming rational behavior. The rate of aircraft procurement is bounded by aircraft production constraints. Formulation V is an adjusted version of formulation IV. Formulation V has two extra variables for each aircraft type, which are bounded by the aircraft production and salvage rates. The maximum salvage rate is determined by the distribution of aircraft service life for each aircraft type. The constraints considered by Formulation V are:

1. Airport capacity: The annual number of operations at an airport cannot exceed its annual service volume.
2. Fleet utilization: The total number of aircraft hours used in the schedule cannot exceed the available aircraft hours for each aircraft type.
3. Runway length: Heavier aircraft cannot take-off or land in airports that have short runways.
4. Airline market niche: Segments with lower demand, in general, are not served by the mainline carriers. Therefore, markets with low demand cannot be served by larger aircraft used by mainline carriers.
5. Payload range constraints: Small aircrafts cannot be used to service segments with relatively larger stage lengths.

6. Aircraft type feasibility: The total frequency reduction in a segment for each aircraft type cannot be greater the service frequency of the aircraft type on the segment.
7. Maximum and minimum segment load factor: The seat capacity offered on the segment should not exceed the segment demand times the reciprocal of the minimum allowable load factor, and should not be less than the segment demand times the reciprocal of the maximum allowable load factor.
8. Minimum and maximum segment frequency: The segment frequency cannot be less than a certain minimum frequency and cannot exceed a maximum allowable value.
9. Allowable frequency variations: The variations in segment frequency between two successive time periods cannot lie outside an allowable bound.
10. Flow conservation: The number of aircraft departing and arriving at an airport should be equal
11. Maximum aircraft production rate: The number of aircraft procured per year cannot exceed the maximum possible annual production capacity for the aircraft type.
12. Maximum aircraft salvage rate: The number of aircraft salvaged per year cannot exceed the maximum possible salvage rate of the aircraft type.

Mathematically Formulation V can be expressed as

Minimize

$$\sum_i \sum_j \sum_k C_k T_{ij} \Delta N_{ijk} + \sum_k (Y_k y_k - Z_k z_k) \quad (320)$$

subject to:

$$(305) - (319)$$

$$0 \leq y_k \leq (y_k)_{\max} \quad (321)$$

$$0 \leq z_k \leq (z_k)_{\max} \quad (322)$$

The constants in equations (307) – (313) are derived from the properties of the current operating schedule in the T-100.

4.3.1.1.7) Formulation VI

Formulation V attempted to generate a new schedule by incrementally changing the schedule of the previous year. Formulation VI attempted to minimize the change in cost, with an upper and lower bound on the allowable change in frequency. Formulation V produced results that were closest to the actual T-100 operating schedule. However, Formulation V suffered from the same drawback as Formulation I, namely underutilization of regional jets. The algorithm replaced regional jets with larger aircraft like the Embraer 170 and the Boeing 737 in segments with significant demand and with turboprops, e.g., Beechcraft 1900 on segments with low demand. The algorithm also scaled down the number of operations on segments in order to accommodate the increased rate of use of the larger jets and turboprops. The reason for the underutilization

of regional jets is due to the high cost per seat of the regional jets. The cost per seat per hour for each of the aircraft types is shown in Table 4.5

The table illustrates that regional jets have the highest cost per seat per hour of the 10 aircraft types. Therefore, the objective function for Formulation V in an attempt to minimize the operating cost does not utilize regional jets until all the other aircraft types have been fully utilized. However, cost minimization is just one of the objectives of an air carrier, the objective being revenue maximization. Since revenue is an increasing function of service frequency, airlines often maximize frequency by using expensive regional jets. The objective function is a combination of cost minimization and frequency maximization. Formulation VI attempts to achieve both objectives at the simultaneously. The weight for each objective function could be varied, depending on the relative importance of both objectives. Formulation VI is similar to Formulation III. A lower weight for the cost minimization leads to most or all the regional jets being used, however the error in the solution relative to the T-100 is larger. A larger weight for the cost minimization leads to better solution accuracy but underutilizes the regional jets.

Table 4.5 Cost per Seat per Hour by Aircraft Type.

Aircraft Cluster	Cost Per Seat Per Hour
1	33.78
2	28.47
3	35.47
4	28.40
5	23.74
6	19.22
7	39.64
8	22.75
9	22.47
10	25.25

Mathematically, formulation VI can be written as

$$\text{Maximize} \quad \sum_i \sum_j \sum_k \Delta N_{ijk} \quad (323)$$

$$\text{subject to:} \quad \sum_i \sum_j \sum_k C_k T_{ij} \Delta N_{ijk} + \sum_k (Y_k y_k - Z_k z_k) \leq (1 \pm \Delta) \pi^* \quad (324)$$

(305) – (319), (321) – (322)

Where, π^* is the minimum operating cost derived from Formulation V. However Formulation VI still suffered from an unacceptable aircraft utilization and deviation from the base schedule. The constants in equations (307) – (313) are derived from the properties of the current operating schedule in the T-100.

4.3.1.1.8) Formulation VII

Formulations V-VI captured the changing fleet over time, However formulations V-VI suffered from the same drawbacks as Formulations I-V, and Formulation V heavily underutilized the smaller aircraft types since it attempted to minimize operating costs. Formulation VI attempted to find a balance between minimizing cost and maximizing frequency. However, the formulation still resulted in about 60% of Regional Jet underutilization. Formulation VII attempts to minimize the aircraft acquisition costs alone without optimizing the aircraft operating costs. Without the fleet change term this formulation would be the same as finding the basic feasible solution to Formulations I-IV. Mathematically Formulation VII can be written as,

Minimize

$$\sum_k (Y_k y_k - Z_k z_k) \quad (325)$$

subject to:

(305) – (319), (321) – (322)

Formulation VII resulted in a schedule that had the minimum deviation from the existing base schedule and also had acceptable aircraft utilization. Formulation VII was chosen as the final formulation due to its accuracy and flexibility. A comparison of all the different formulations is illustrated in Section 5.2. The constants in equations (293) – (298) are derived from the properties of the current operating schedule in the T-100.

4.3.3.2) Integer Solution Generator

The optimal frequencies obtained from the frequency generator models are the solutions to the LP relaxations of the original integer programs. Therefore, in general these frequencies are real numbers. However, aircraft frequencies need to be integers. The integer constrained frequency generator is a very large-scale integer program (26,000 variables and 107,000 constraints) and deriving an exact integer optimal solution is an arduous task, Therefore the integer solution generator obtains an integer solution from the continuous solution to the frequency generator by a rounding scheme based on a Lagrangian dual formulation of the original primal problem. Sherali (2008) suggested that an intelligent rounding scheme that minimized the degradation in the objective of the LP relaxation could be used to derive an acceptable integer approximation of the LP

relaxation. To present this procedure, consider, Formulation VII can be written in generic matrix form as follows:

Minimize

$$\sum_k (Y_k y_k - Z_k z_k) \quad (326)$$

subject to:

$$\sum_j \lfloor N_{ijk} - \overline{N_{ijk}} \rfloor - \sum_j \lfloor N_{jik} - \overline{N_{jik}} \rfloor = 0, \forall i, k \quad (327)$$

$$\sum_k A_k N^k + \sum_k \alpha_k y_k + \sum_k \beta_k z_k \geq b \quad (328)$$

$$N^k \geq 0, \forall k, \quad (329)$$

where for each k , N^k is a vector of variables $(N_{ijk}, \forall i, j)$ with $N_{ijk} = \overline{N_{ijk}} + \Delta N_{ijk} \forall i, j, k$; Equation (327) represents the equality constraint (318) and the remaining constraints in (305)-(317) written compactly as (328) for appropriate matrices A_k and vectors α_k, β_k and $b, \forall k$. Let $N_{ijk}^*, \forall i, j, k; y_k^*, \forall k; z_k^*, \forall k$ solve the linear relaxation of Formulation VII with an optimal dual variable $\pi^* \geq 0$ associated with Constraint (298). Then, the Lagrangian dual function for this LP corresponding to dualizing (298) using the dual optimal solution π^* is given by

$$\pi^{*T} b - \sum_k \pi^{*T} A_k N^k, \quad (330)$$

where by LP duality, the reduced costs for the y_k and z_k variables are zeros (see Bazaara et.al, 2005). Motivated by the fact that minimizing (330) subject to (327) and (328) would yield the same LP value as solving (326)-(329) and noting that the resulting LP is separable over the aircraft types k , the following rounding *integer solution generator* problem (**ISG_k**) for each k is formulated:

ISG_k: Minimize

$$- \pi^{*T} A_k N^k \quad (331)$$

subject to:

$$\max\{0, \lfloor N_{ijk}^* - \Delta_k \rfloor\} \leq N_{ijk} \leq \lfloor N_{ijk}^* + \Delta_k \rfloor, \forall i, j, k \quad (332)$$

$$\sum_j \lfloor N_{ijk} - \overline{N_{ijk}} \rfloor - \sum_j \lfloor N_{jik} - \overline{N_{jik}} \rfloor = 0, \forall i, k \quad (333)$$

$$N_{ijk} = 0 \text{ if } N_{ijk}^* = 0, \forall i, j \quad (334)$$

$$N_{ijk} \geq 0 \text{ and integer, } \forall i, j, k, \quad (335)$$

where Δ_k is an allowable manipulation in the rounded down value $\lfloor N_{ijk}^* \rfloor$ for obtaining N_{ijk} , and is adjusted to be as small as possible to yield feasibility for ISG_k, $\forall k$. It should be noted that because of the dualized constraints (328), this rounding

mechanism need not produce a feasible solution to Formulation VII (except for the retained constraints (334) and (335)), but the method explicitly provides a practical alternative to derive an “acceptable” integer solution by way of rounding the LP optimal solution.

To compare the solutions produced by ISG_k, $\forall k$ with Formulation VII, the original formulation was solved using commercial software package CPLEX11.0. The results of the comparison are illustrated in Section 5.2.2.1.

4.3.3.3) Daily Frequency Generator

The daily frequency converts the yearly integer frequency values obtained from the integer solution generator to an equivalent daily frequency. The minimum threshold for regular operations is taken as 250 flights per year (5 flights per week). Therefore, the yearly frequency can only be a linear combination of flights that operate 5 days, 6 days, and 7 days per week (250 flights per year, 307 flights per year and 365 flights per year). The daily frequency generator can be expressed mathematically as:

Maximize

$$250 X_{ij} + 307 Y_{ij} + 365 Z_{ij} \quad (336)$$

subject to:

$$250X_{ij} + 307Y_{ij} + 365Z_{ij} \leq N_{ij}^*, \forall i, j \quad (337)$$

X_{ij}, Y_{ij}, Z_{ij} integer

Where for each (i, j) :

X_{ij} is the daily frequency of flights that operate 5 times per week.

Y_{ij} is the daily frequency of flights that operate 6 times per week.

Z_{ij} is the daily frequency of flights that operate daily.

N_{ij}^* is the yearly frequency from the origin airport (i) to the destination airport (j) derived from the integer solution generator.

4.3.3.4) Schedule Generator

The schedule generator obtains the operating daily schedule from the output of the daily frequency generator. The schedule generator assigns a flight number, airline ID, departure and arrival times and aircraft class to each flight obtained from the daily frequency generator. The objective of the flight schedule is to minimize the passenger schedule delay. Teodorovic (1983) develops a model to determine the flight schedule that minimizes the passenger schedule delay. The total daily demand is given by

$$H(t) = \int_0^t h(t) dt \quad (338)$$

Where:

$H(t)$ is the total number of passengers that desire to travel up to time t .

$h(t)$ is the probability density function of the travel demand during the time interval t .

Teodorovic (1983) assumes that a passenger would adjust his departure time to take the nearest flight. Let $x_1, x_2, \dots, x_N$ be the departure times of N flights that depart an airport on during time period T . The total delay accrued by passengers over the time period is given by

$$W^* = 2 \int_0^{x_1} H(t)dt + 2 \int_0^{x_2} H(t)dt + \dots + 2 \int_0^{x_N} H(t)dt - 2 \int_0^{(x_1+x_2)/2} H(t)dt - 2 \int_0^{(x_2+x_3)/2} H(t)dt - \dots - 2 \int_0^{(x_{N-1}+x_N)/2} H(t)dt - x_N H(T) - \int_0^T H(t)dt + TH(T) \quad (339)$$

To minimize the passenger waiting time,

$$\frac{\partial W^*}{\partial x_i} = 0 \quad (340)$$

By evaluating the partial derivatives and setting the derivatives to zero, we get

$$H(x_1) = \frac{H\left(\frac{x_1 + x_2}{2}\right)}{2} \quad (341)$$

$$H(x_2) = \frac{H\left(\frac{x_1 + x_2}{2}\right) + H\left(\frac{x_2 + x_3}{2}\right)}{2} \quad (342)$$

$$H(x_N) = \frac{H\left(\frac{x_{N-1} + x_N}{2}\right) + H(T)}{2} \quad (343)$$

By setting the derivatives to zero, Teodorovic (1983) obtains a series on non-linear simultaneous equations, which can then be solved for the departure times. The author however notes that the model requires the distribution of the passenger desired departure time in each market. In this paper, we assume that the flight departure times for the baseline schedule as the desired passenger departure times and the schedule for the next year is based on the scheduled departure times for the baseline year.

This approach does not take into account the possibility of the hourly demand exceeding the hourly airport capacity, especially during peak times. In order to prevent airport overcapacity, the schedule needs to be smoothed when the demand exceeds capacity.

4.3.3.5) Scheduling of Direct Stopover Flights

The airline schedule has a significant proportion of direct stopover flights. A stopover flight is defined as a flight with the same aircraft and same flight number that serves multiple airports. An example is the Southwest flight from Birmingham, AL (BHM) to Ford Lauderdale, FL (FLL). In 2004 there was no direct service from BHM to FLL. However, Southwest flight 507 had a service from BHM to FLL via Orlando, FL (MCO). It is necessary to account for stopover flights as a significant proportion (about 16%) of the total flight schedule in 2004. Neglecting stopover flights could lead to significant

errors in the network topology as some airports pairs are served only by direct stopover flights. For example, in 2004 the BHM-FLL segment was only served by Southwest flight 507. The demand estimation for stopover flights is not straightforward as the demand is the combined demand for the constituent segments. The segment demand for the BHM-MCO segment is the sum of the demand from BHM to MCO and the demand from BHM to FLL. The BHM-MCO-FLL flight cannot be scheduled independently as its schedule depends on the schedule of the BHM-MCO and MCO-FLL segments. In the model the stopover flights are scheduled using the best combination of the non-stop flights of the individual constituent segments. For example, for the BHM-MCO-FLL flight, the combination of BHM-MCO and MCO-FLL flights that give the minimum stopover time at MCO is chosen as the schedule for the BHM-MCO-FLL stopover flight. The model also ensures that the constituent flights have the same aircraft type and same airline. The model assumes that each stopover flight has unit frequency in order to ensure that sum of the frequency of stopover flights from an airport does not exceed the frequency of the non-stop flights.

4.4) Network Evolver Model

The network evolver was introduced briefly in Section 3.3.4. The network evolver changes the structure of the airline network by adding and removing routes. Airlines add and delete routes based on economic and financial considerations. Airlines also add point-to-point and hub-to-spoke connections based on airport-to-airport distance and demand. Airlines could also start service from new secondary airports to relieve congestion at primary airports. The network evolver model is composed of five sub-modules

- (1) Cutting service on non-profitable routes
- (2) Adding new point-to-point service
- (3) Adding new hub-to-spoke connections
- (4) Adding and removing hubs
- (5) Introducing service from airports that have no significant commercial service at present

4.4.1) Cutting Service on Non-profitable Routes

Airlines frequently cull routes that do not produce adequate profit, do not have a high enough load factor, or do not have a high enough base frequency. For example in 2004, US Airways removed many connections to Pittsburgh International Airport (PIT) and downgraded the airport from a hub to a focus city. Following the de-hubbing of Pittsburgh International Airport, the number of enplanements in Pittsburgh dropped by about 50%. These effects have to be captured in order to obtain a good picture of the airline industry. However, some airports have service despite low load factors and significant losses. These airports are called Essential Air Service Airports and are subsidized by an act of the U.S Congress. These airports therefore, are treated as exceptions. The model derives the likelihood of the cessation of a route based on certain explanatory variables. The likelihood is then converted into a binary decision, by the

application of a threshold variable. The independent variables considered in the model are

- (1) Profit/Seat – As was explained in the last paragraph airlines cut service in routes that have poor revenue or profit potential. Consequently the profit per seat earned by airline is one of important variables that affect an airline's strategy
- (2) Load Factor – An airline attempts to fill as many seats as possible in each flight. Any flight that is mostly devoid of passengers has no revenue potential and it is highly likely that the airline would cease to operate the flight.
- (3) Distance – Automobiles and transit compete for market share with airlines on short-haul routes. Therefore, the revenue potential for a short-haul route is less than that of medium and long haul routes and is much more unattractive for revenue generation. An airline facing losses would cease service in short-haul routes before it culls medium or long haul routes.
- (4) Frequency – The probability of service being ceased on a route is dependent on the density of service on the route. The chances of an airline ceasing service on a high-density route between two major cities are very low. On the other hand, low-yield routes, which primarily serve small communities, are routes, which are cut by the airlines, when facing financial difficulties
- (5) Essential Air Service airports – The Essential Air Service United States Government program that was enacted to guarantee that small communities in the United States that prior to deregulation were served by certificated airlines maintained commercial service(Wikipedia(2006)). It was designed to maintain a minimal level of scheduled air service to these communities which otherwise would not be profitable. The United States Department of Transportation currently subsidizes airlines to serve approximately 140 rural communities across the country that otherwise would not receive any scheduled air service. Therefore, an air carrier would continue service on routes, which are otherwise undesirable from an economic, and operation point-of-view and can be described as “outliers” as these routes are subsidized by the federal government. The Essential Air Service indicator variable is introduced to remove the bias in the model, which arises due to the existence of these subsidized routes.

4.4.1.1) Profit Computation Sub-module

The profit for each route is one of the most important variables that affect the airline strategy. The profit was computed by subtracting the revenue and the cost. The cost of the routes was obtained from the operating cost of the aircraft flown on the route. The operating costs of a few of the sample aircraft is given in Table 4.5. The costs do not include indirect operating costs such as salaries of non-flying personnel such as CEO salaries and non-operating costs such as advertising costs and cost of the ticket counters. The operating costs derived from the model were converted to the total costs by a utilizing a multiplication factor. The multiplication factor was the ratio of the total costs to the operating costs in the year 2004. The total operating costs were obtained from the Form P-6. The multiplication factor was assumed to be constant for the future years.

The operating cost of the segment was obtained by multiplying the operating cost per hour by the number of aircraft hours spent by each aircraft type flying the route. The operating cost of a few sample aircraft is given in Table 4.6

Table 4.6 Aircraft Operating Costs from the T-100.

AIRCRAFT OPERATING EXPENSES (X1,000 DOLLARS)	AIR TIME (MINUTES)	OPERATING COST PER HOUR	AIRCRAFT NAME
13,225	2,797,559	284	Cessna C-402/402a
17,721	1,219,331	872	British Aerospace Jetstream 31
2,654	59,143	2,692	Aerospatiale/Aeritalia Atr-42
46,047	1,719,655	1,607	Dornier 328
56,159	2,337,439	1,442	Dehavilland Dhc8-300 Dash 8
75,497	2,726,326	1,662	Dehavilland Dhc8-200q Dash-8
118,205	5,981,273	1,186	Dornier 328 Jet
482,948	23,032,558	1,258	Saab-Fairchild 340/B
2,073,716	75,567,183	1,647	Embraer-145
173,265	5,905,877	1,760	Avroliner Rj85
1,205	24,433	2,958	Mcdonnell Douglas Dc-9- 15f
2,505	46,222	3,251	Boeing 727-100
92,422	1,720,263	3,224	Mcdonnell Douglas Dc-9- 40
1,945,768	52,851,797	2,209	Boeing 737-700/700lr
7,251	99,604	4,368	Mcdonnell Douglas Dc9 Super 87
314,533	2,822,495	6,686	Boeing 777

The revenue computation process was not straightforward as the ticket revenue from passengers is obtained for the whole route, unlike costs which are defined for each segment. The route revenue therefore had to split between individual legs. A few schemes for splitting the revenue were considered.

- (1) Segment Distance
- (2) Segment Operating Costs
- (3) Average Segment Ticket price

The scheme that was finally chosen was the average segment ticket price. Segment distance was the simplest scheme to split the route revenue between its constituent

segments. However, the scheme assumes a linear relationship between cost and flown distance, which is not true as smaller aircraft are more expensive on a cost per seat mile basis than larger aircraft. The second scheme is more accurate, however the second scheme would derive a different split between the segments depending on the aircraft type, which is not very intuitive as revenue is not directly a function of aircraft type. The second scheme was abandoned and third scheme was adopted. The revenue computation process is illustrated in detail in Figure 4.5.



Figure 4.5 Sample Computation of Revenue between Roanoke,VA(ROA) and Atlanta,GA(ATL).

Assuming that the ROA to SEA market has a revenue of \$ 100,000, the ROA to SFO market revenue \$150,000 and the ROA to LAX market is \$200,000.

From Figure 4.5 it can be inferred that the revenue for the ROA→ATL leg is,
 $(154/327)*100,000 + (154/346)*150,000 + (154/328)*200,000 = \$ 207,760$.

There were a considerable number of stopover segments in the network. These segments had to be modeled separately as, the above scheme assumes that the segments are non-stop. The stopover segment revenue was computed in a two-stage process. In the first stage the revenue for the stopover segment was derived by the scheme described in the previous paragraph. The stopover segment revenue was then divided into its constituent non-stop segments by applying the same scheme once again on the stopover segment. The final revenue values were only computed for non-stop segments. The model assumes that removal of a non-stop segment would imply that all the stopover segments which utilize this non-stop segment would be removed as well.

4.4.2) Introduction of New Point-to-Point Service

Airlines add point-to-point service often bypassing major hubs, when the demand between the two airports becomes large enough to economically provide direct service. For example, if a traveler desires to fly from Roanoke, VA to San Francisco, CA, the airlines do not offer direct service between the two cities. The traveler has to connect at one of the hubs operated by the airlines, to fly to San Francisco. This is because the distance between Roanoke and San Francisco is large and the demand between the two cities is too small to justify direct service. The model adds direct service between two points when the demand between two airports is greater than a certain threshold. The threshold is dependent on distance as a greater distance between airports would imply that a greater demand must exist between the airports to introduce direct service. The threshold is computed by calculating the demand which would allow an airline to operate a minimum frequency of 250 flights per year (5 days per week) at 70% load factor. The typical operating aircraft becomes larger with increasing distance, thus the impedance for direct service increases with distance. The threshold values for the introduction of point-to-point service are given in Table 4.7.

Table 4.7 Distance Threshold for the Introduction of Point-to-Point Air Service.

Distance Range (Miles)	Demand Threshold (Yearly Demand)
0 - 200	1,575
200-641	3,325
641-1000	7,875
1000-1300	12,600
>1300	22,225

4.4.3) Introduction of New Spoke-to-Hub Flights

Airlines add new service to their hubs, if the demand at a single spoke airport becomes significant and the airport is not directly connected to an airline hub. The criteria for adding new hub-spoke routes are different from those that decide new point-to-point routes. This is because most of the passengers that utilize a spoke-hub connection do not stop at the hub but continue their journey to other destinations from the hub. The hub-spoke demand is not a factor to add service between the hubs and spoke. The model assumes that the hub-spoke relationship is symmetric, in other words introduction of a spoke to hub service automatically implies the introduction of the hub to spoke service. The variables considered in the model are

- 1) Distance from the Spoke to Hub: The distance from spoke to the airline hub is an important variable that influences the likelihood of a connection between the spoke to the hub, as larger distances would tend to introduce considerable resistance to the introduction of the service

- 2) Airline hub connectivity of the spoke airport: The connectivity of the spoke airport is another important factor in determining the attractiveness of the spoke-hub connection. If the spoke airport is connected to an existing airline hub, the incentive to connect the spoke to another airline hub decreases
- 3) Maximum frequency from the spoke airport: The maximum frequency of service from the spoke airport is another indication of connectivity of the spoke airport. If the maximum frequency is large, which is an indication of a well-connected spoke airport, the increased utility of another hub connection from the spoke is low
- 4) Total demand from the spoke: A spoke airport with a high demand for air travel would induce an airline to connect it the airport to its hubs. On the other hand, if the airport demand is low, then the airline has little or no incentive to improve the airport connectivity.

The model uses a binary probit regression technique to derive the likelihood of spoke to hub service, from the four independent variables described in the previous paragraph. The likelihood is then converted to a binary response by utilizing a threshold value.

4.4.4) Addition and Removal of Airline Hubs

The Federal Aviation Administration classifies the commercial airport set into four categories namely large, medium, small and non-hubs. A large hub is defined as an airport that handles at least 1 percent of the total passenger enplanements in the United States. A medium hub is defined as an airport that handles between 0.25 and 1 percent of the total US passenger traffic. A small hub handles between 0.05 and 0.25 percent of the total US passenger traffic and a non-hub handles between 10,000 passengers and 0.05 percent of the total US passenger traffic. The model however attempts to dynamically change the set of airline hubs, which could be different from the hub classification designed by the FAA. Most of the airline hubs are large hubs; however some airline hubs such as Cleveland (CLE) and Milwaukee (MKE) are medium hubs. Therefore, it is necessary to derive a definition of an airline hub, which is independent of the FAA definition. In this model, any airport that has a connection to a minimum of 53 other airports and a minimum of 15 small and non-hubs as defined by the FAA is assumed to be an airline hub. Conversely an existing airline hub that fails to meet this criterion due to route closures is removed from the list of airline hubs. An airport that is not an airline hub can add links only through point-to-point connections. However, if enough point-to-point connections are added to an airport which is not a part of the airline hub set, it could meet the criterion for an airline hub. The airport could then grow through both the addition of point-to-point and spoke-to-hub service. This module could capture the emergence of future airline hubs.

4.4.5) Introduction of Service from New Airports

In the year 2006, the continental United States had 410 airports, which had commercial service. However, the US also has about 580 airports with paved runways of length 6,000 feet or greater which could potentially serve as secondary airports when some of the 443 primary airports reach capacity.. The model selects the best candidate for developing commercial service in two stages. In first stage all the secondary airports which are more than 150 miles from the centroid of the Metropolitan Statistical Area (MSA) served by the primary airports are discarded. The analysis only considers MSA's as most of the congested areas are situated in urban areas. In the second stage, the remaining candidates are ranked according to certain criterion and best candidate airport is chosen for commercial service development. The criteria considered are:

- (1) Airport Capacity: The capacity of the reliever airport is the most important factor in the analysis, since the purpose of the reliever airport is to mitigate congestion at the primary airport. Therefore, a secondary airport with a higher capacity would ensure a greater reduction of demand at the primary airport
- (2) Distance from the MSA centroid: The secondary airport should be as close as possible to the MSA population centroid to attract the maximum possible demand. If the reliever airport is too far away from the MSA then it would attract very less demand, and fail to mitigate congestion at the primary airport
- (3) Length of the longest runway: Runway length is an important factor that could constrain the number and nature of operations at an airport. A reliever airport with a long runway could serve heavier aircraft and consequently handle higher demand.
- (4) Presence of control tower: A control tower is essential for realizing the entire capacity potential at an airport. This is because an airport with a tower can offer better traffic management and consequently have a greater capacity than a non-tower airport. A airport with a tower could also enhance the operational safety at an airport

The model assigns a rank for all the candidates based on the criterion described in the previous paragraph and the candidate with the best rank is chosen as the reliever airport.

Chapter 5 Model Results and Validation

5.1) Introduction

The model described in Chapter 4 has four components. These are

- (1) The Air Traffic Network Flow Problem
- (2) The Schedule Generator
- (3) The Network Evolution Sub-Module
- (4) The Fare Evolution Sub-Module

These sub-modules need to be validated separately and as a part of the Transportation System Analysis Model (TSAM). Therefore, the model validation is composed of two parts. In the first part, the individual model results are validated against real data from the Department of Transportation (DOT) such as the T-100 and the DB1B. In the second part the results of the integration of the model with TSAM are validated against real data. The trends derived from executing TSAM into the future years are also plotted to ascertain their plausibility.

5.2) Sub-Module Results and Validation

In this section, the results of the individual sub-modules are presented along with the validation of the sub-modules by comparing their output to actual passenger flow and airline frequency data.

5.2.1) The Air Traffic Network Flow Problem

As explained in Section 4.1 the air traffic network flow problem attempts to derive route flows between all pairs of origin-destination pairs in the United States. The model is a multinomial logit that predicts the likelihood of a route being chosen based on certain independent route related attributes. The model is validated using statistical techniques such as the likelihood ratio, R^2 and the T-statistic. The route flow, segment flow and the enplanements estimates derived from the model are validated against real route flow and enplanement data from the DB1B and the real segment flow data from the T-100. The result of the air traffic network flow problem and its statistical accuracy is given in Table 5.1. A comparison of the multinomial and nested logit formulations is given in Table 5.2

Table 5.1 Multinomial Logit Results and T-Statistics With the Average Airport Delay for the year 2000.

Variables	Multinomial Logit Coefficients for 2000	
	Direct + Connecting	Pure Connecting
Travel Time	-4.8949 (-324.72)	-4.4508 (-358.79)
Fares	-0.8128 (-162.93)	-1.2543 (-288.15)
Frequency	0.0343 (94.87)	0.0927 (199.05)
Seats	2.0297×10^{-6} (186.69)	3.0347×10^{-6} (236.47)
Legs	-1.4536 (-218.29)	N/A
ATL Dummy	0.2656 (-23.20)	0.3092 (-54.05)
IAD Dummy	-0.4451 (35.11)	-0.6098 (83.88)
L/R Ratio	0.6635	0.1702
R ²	0.9918	0.4547

Table 5.2 Multinomial and Nested Logit Results and T-Statistics with the Average Airport Delay for the year 2000.

Variables	Logit Coefficients for Direct + Connecting in 2000	
	Multinomial	Nested
Travel Time	-4.8949 (-324.72)	-5.1052 (-319.76)
Fares	-0.8128 (-162.93)	-0.8361 (-163.68)
Frequency	0.0343 (94.87)	0.0339 (91.75)
Seats	2.0297×10^{-6} (186.69)	2.0864×10^{-6} (216.07)
Legs	-1.4536 (-218.29)	-1.4309 (-211.87)
ATL Dummy	0.2656 (35.11)	0.0400 (2.80)
IAD Dummy	-0.4451 (-23.20)	-0.4357 (-22.67)
INC1	N/A	0.8453 (295.28)
INC2	N/A	0.8714 (329.83)
L/R Ratio	0.6635	0.664
R ²	0.9918	0.9918

From Table 5.1, it can be observed that the coefficient for both travel time and cost are negative. This result is expected as a higher travel time or cost would result in a higher disutility for the traveler and consequently a lower chance of a traveler taking the route. The coefficient for frequency is positive, as a higher frequency implies a lower schedule delay and a greater attractiveness for the route. The coefficient for the number of seats is positive as the probability of a traveler being denied a seat in his preferred itinerary is

inversely proportional to the number of seats offered by an airline on the route. From Table 5.2, it can also be observed that a nested logit model does not lead to any significant increase in model accuracy.

The above calibration was performed for the year 2000. However, the airline network has undergone significant changes in cost, schedule and connectivity especially after September 11, 2001. The model was re-calibrated for the year 2006 and the results of the calibration are shown in Tables 5.3 and 5.4

Table 5.3 Multinomial Logit Coefficients for the Year 2006.

Variables	Multinomial Logit Coefficients for 2006	
	Direct + Connecting	Pure Connecting
Travel Time	-4.6971 (-323.22)	-5.6296 (-653.07)
Fares	0.1828 (45.73)	-0.4706 (146.83)
Frequency	0.0714 (186.57)	0.1414 (305.83)
Seats	2.0937×10^{-6} (198.38)	2.3834×10^{-6} (187.19)
Legs	-1.8527 (-218.29)	N/A
L/R Ratio	0.7129	0.1831
R ²	0.9327	0.5037

From Table 5.3 it can be seen that the fare variable is positive which is counter-intuitive. Therefore, the model was re-run with the fare variable dropped.

Table 5.4 Multinomial Logit Model for the Year 2006 with the Fare Variable Dropped.

Variables	Multinomial Logit Coefficients for 2006	
	Direct + Connecting	Pure Connecting
Travel Time	-4.5943 (-319.72)	-5.5233 (-644.09)
Frequency	0.0712 (186.06)	0.1430 (309.50)
Seats	2.0778×10^{-6} (197.06)	2.3354×10^{-6} (183.62)
Legs	-1.8819 (-289.27)	N/A
L/R Ratio	0.7126	0.1781
R ²	0.9326	0.4953

The positive value of the fare coefficient in Table 5.3 could be due to high correlation between the number of connections and the fare, especially for markets which have direct flights. To test this hypothesis, the logit model for the O-D pairs which have both direct and connecting itineraries were re-calibrated without the travel time and the variable for the number of legs. The results for the calibration are given in Table 5.5. From Table 5.5

it can be observed that the removal of the variables for travel time and number of legs restores the correct sign for the fare variable. Therefore, it can be concluded that inclusion of both the travel time and fare variable for the 2006 calibration is not a statistically sound calibration. It can also be hypothesized that travel time and fares have become much more correlated in 2006 compared to 2000 as the coefficient for fare was negative in 2000 even with the inclusion of the travel time variable.

Table 5.5 Multinomial Logit Coefficients for the Direct and Connecting Case With the Travel Time and Number of Legs Variable Dropped.

Variables	Multinomial Logit Coefficients for 2006
	Direct + Connecting
Fares	-0.9462 (-902.06)
Frequency	0.0827 (928.58)
Seats	1.3798×10^{-6} (637.24)
L/R Ratio	0.1536
R ²	0.4996

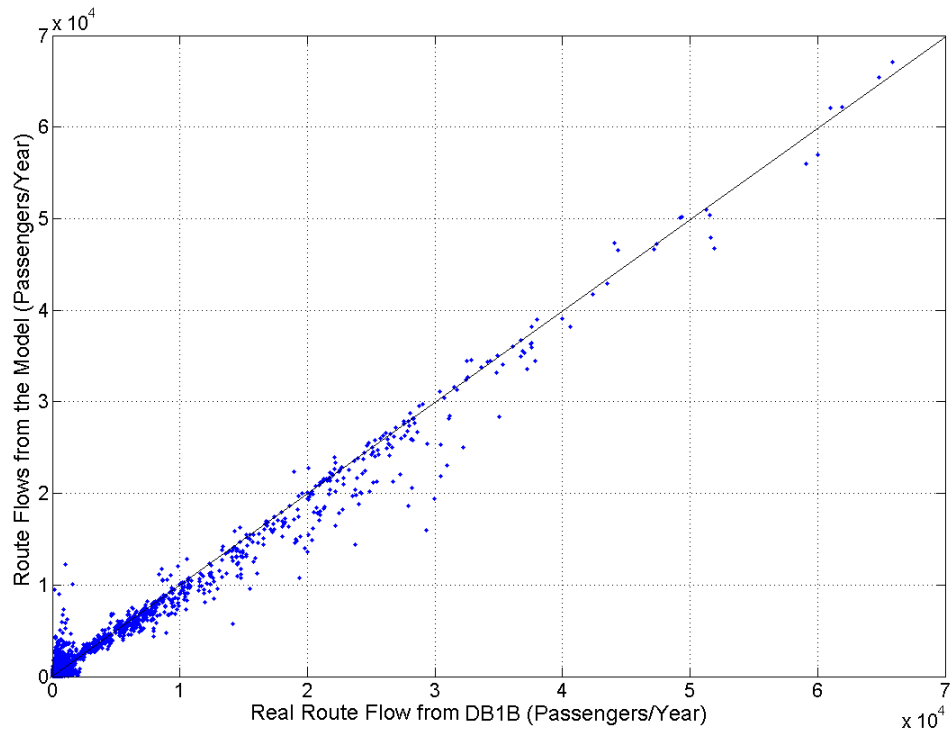


Figure 5.1 Observed vs. Estimated Route Flows for Trips Originating and Terminating at the Large Hubs.

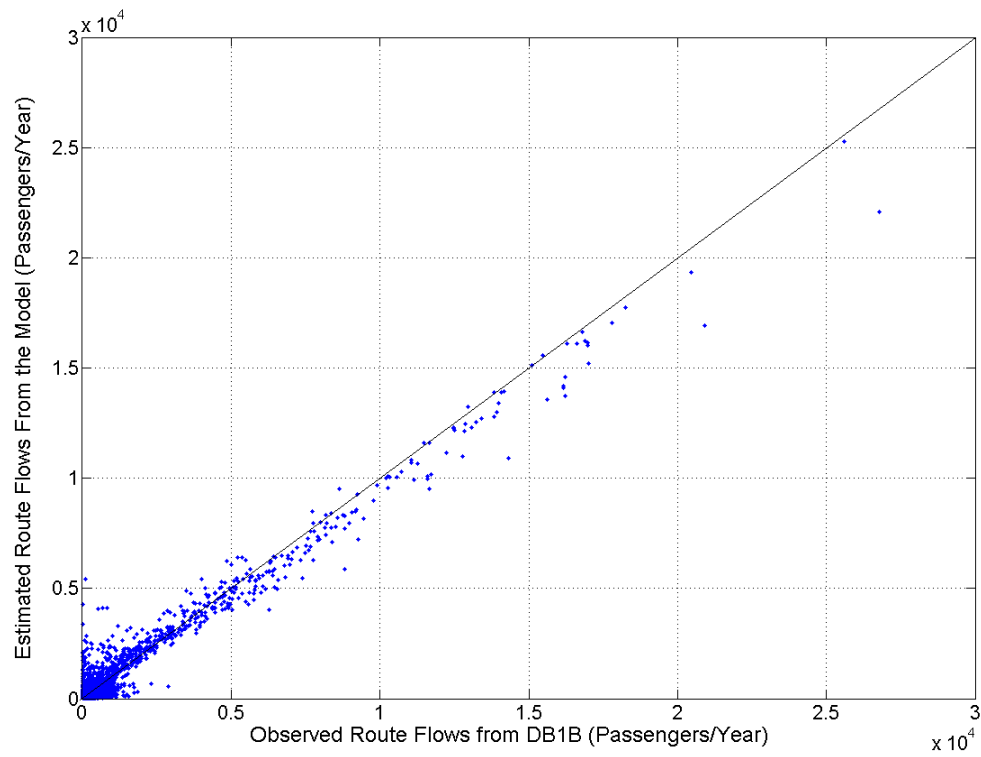


Figure 5.2 Observed vs. Estimated Route Flows for Trips Originating at the Large Hubs and Terminating at the Medium Hubs.

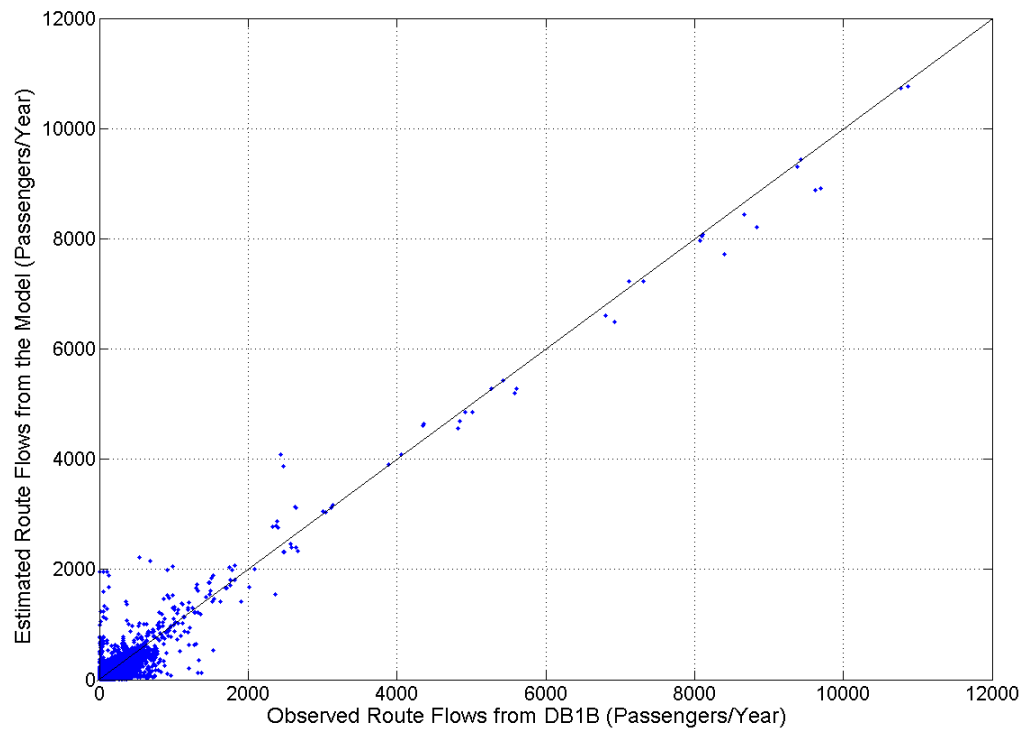


Figure 5.3 Observed vs. Estimated Route Flows for Trips Originating and Terminating at the Medium Hubs.

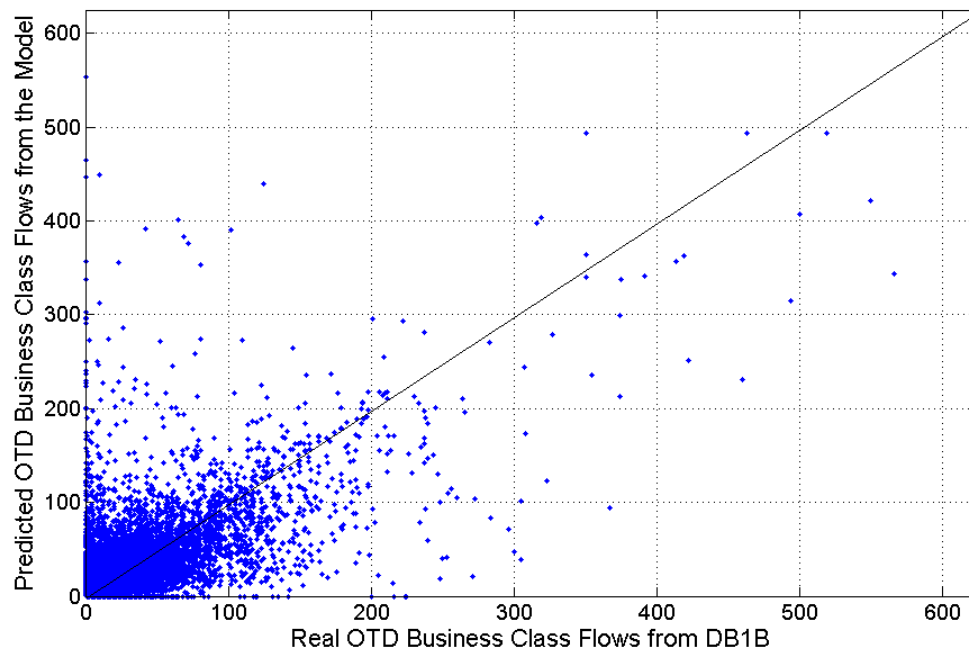


Figure 5.4 Observed vs. Estimated Route Flows for trips Originating and Terminating at the Small Hubs.

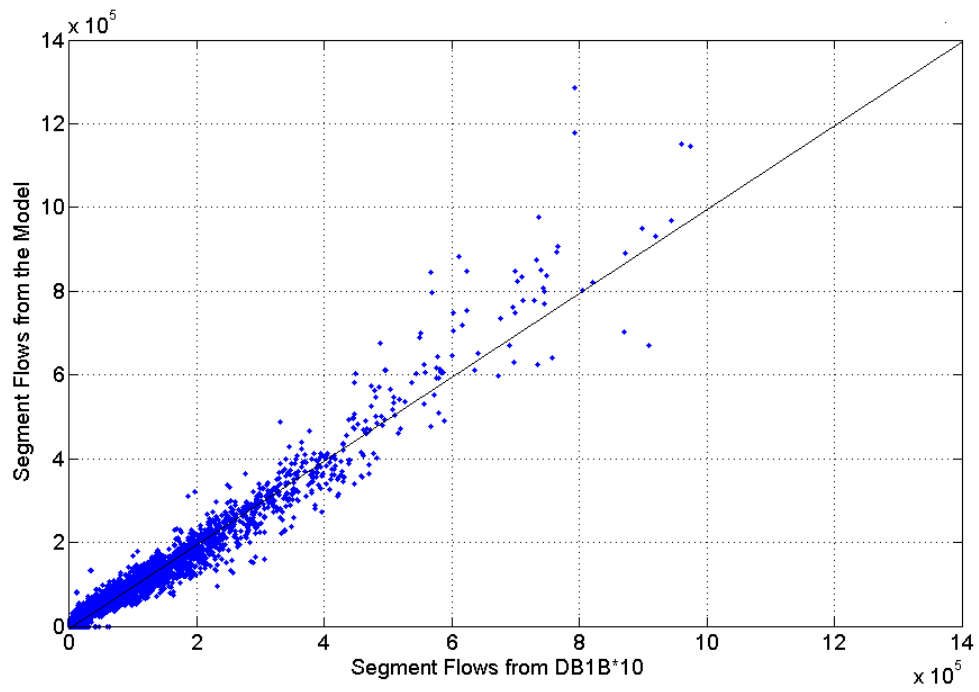


Figure 5.5 Observed and Estimated Segment Flows for the year 2000.

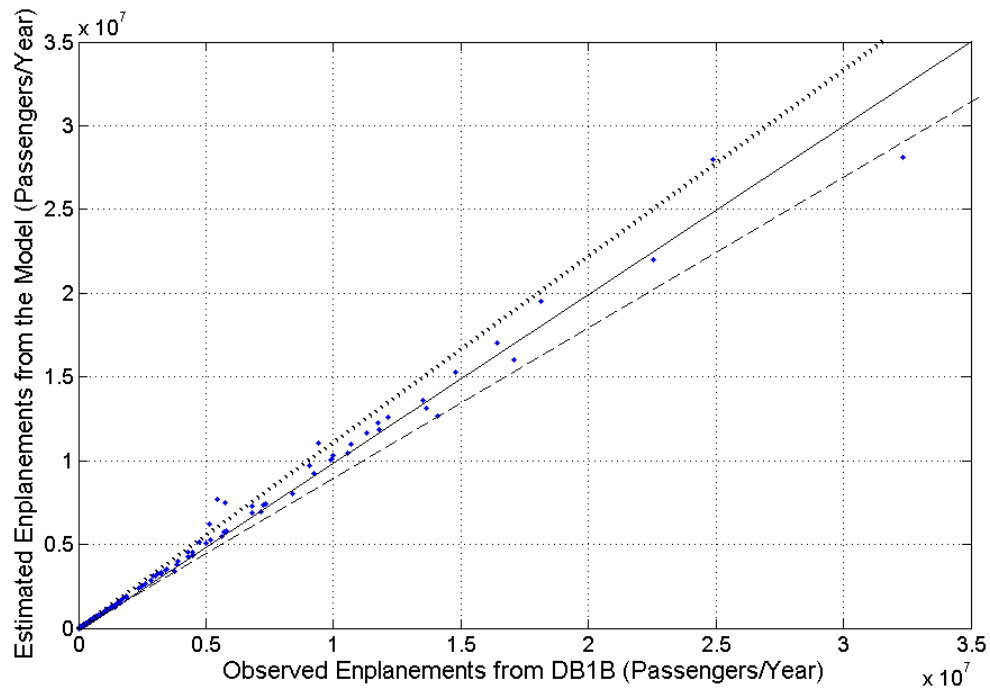


Figure 5.6 Observed Enplanements from the DB1B vs. Estimated Enplanements from the Model Without the Incorporation of Dummy Variables for the Year 2000.

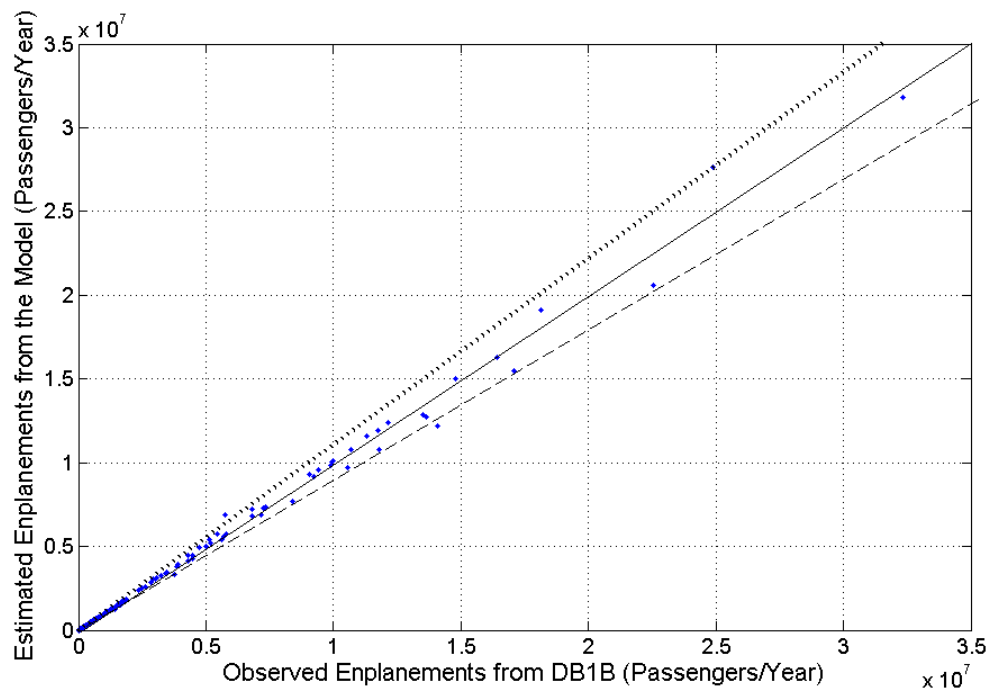


Figure 5.7 Observed Enplanements from the DB1B vs. Estimated Enplanements from the Model With the Incorporation of Dummy Variables for the Year 2000.

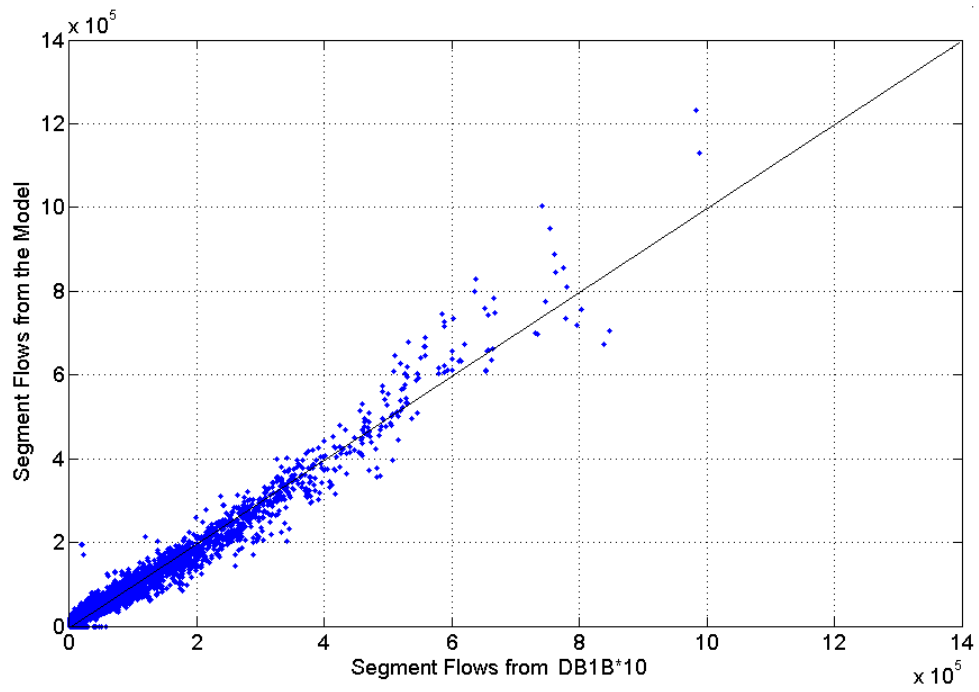


Figure 5.8 Observed and Estimated Segment Flows for the year 2006.

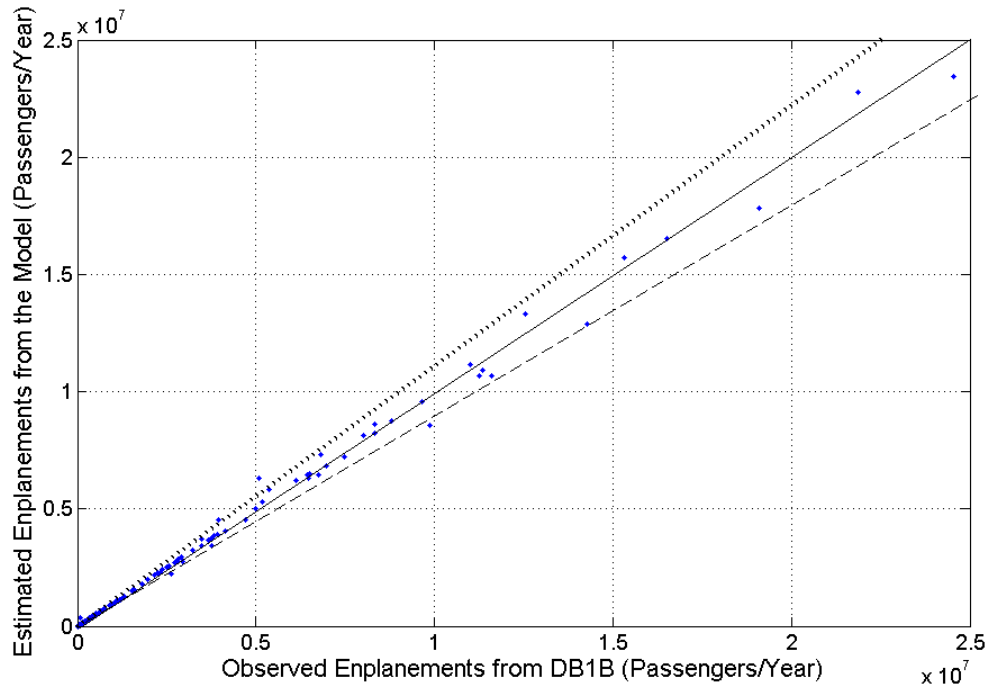


Figure 5.9 Observed Enplanements from the DB1B vs. Estimated Enplanements from the Model Without the Incorporation of Dummy Variables for the Year 2006.

5.2.2) The Schedule Generator

As explained in Section 4.4, several formulations were attempted for solving the schedule generator problem. The accuracy of the different formulations is presented in Figures 5.10 to 5.16. The formulation that has the minimum deviation with the actual operating schedule from the T-100 and with the best fleet utilization is chosen as the final formulation. Formulation VII has the best performance as measured by the sum-of-squared error and is utilized to construct the schedule for each time period.

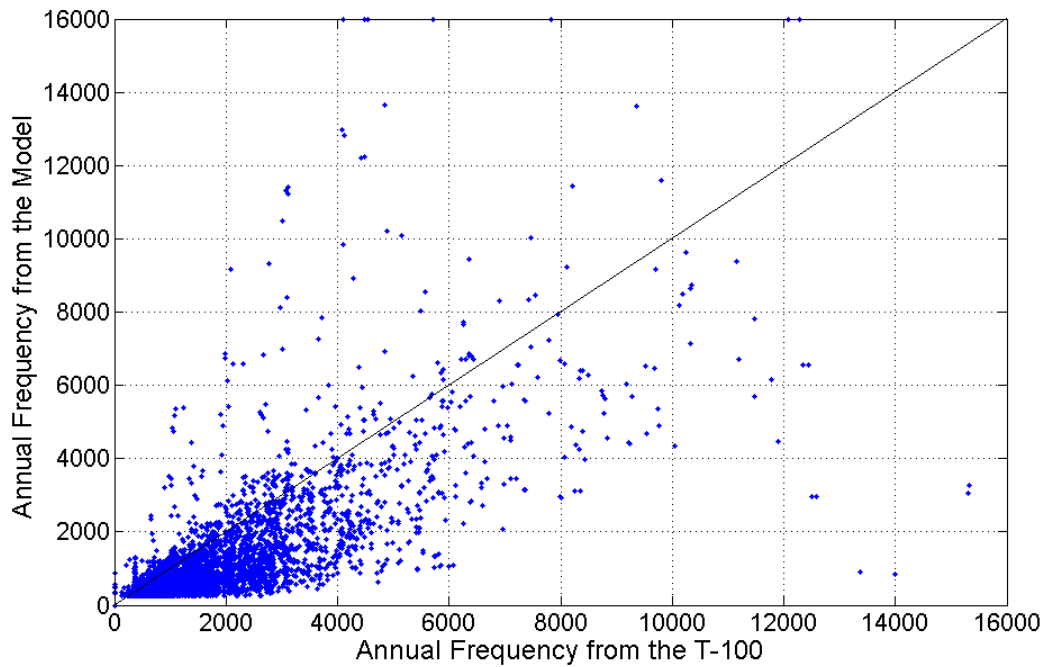


Figure 5.10 Calculated Frequencies from Formulation I vs. Actual Operating Schedule from the T-100.

It can be concluded from Figure 5.10, that Formulation I, is not a good representation of the actual operating schedule from the T-100. The sum-of-squared error for Formulation I, is about 8.5×10^9 , which is a significantly large error. The average error of prediction in a single flight segment, is about 1400 flights per year or about four flights per day. This error is a large error and consequently Formulation I is not an acceptable model of the existing flight schedule in the NAS. The main reason behind the large error associated with Formulation I, is that it focuses only on cost minimization, which leads to a severe underutilization of Regional Jet (RJ) service. This is supported by the fact that only about 5% of the total RJ capacity is utilized. However, airline schedules are not driven by cost minimization alone. Studies by Hansen (1990) and Hong and Parker (1992) indicate that the airline industry is fiercely competitive and the players in a market schedule flights to capture the maximum possible market share, which cannot be captured in Formulation I.

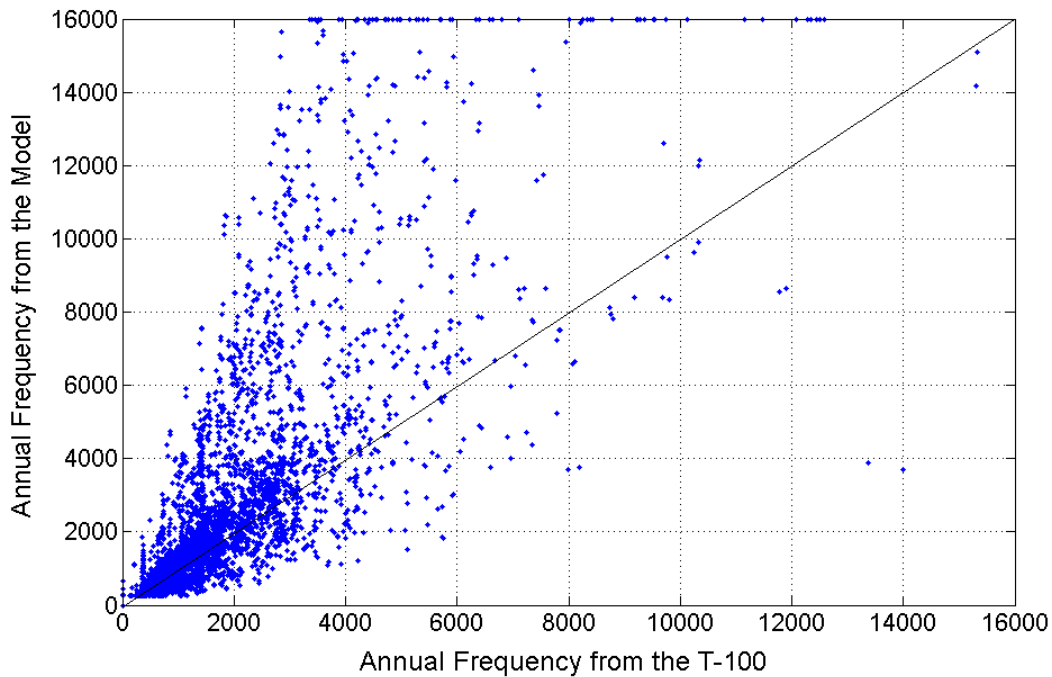


Figure 5.11 Calculated Frequencies from Formulation II vs. Actual Operating Schedule from the T-100.

Formulation I was based on cost minimization which resulted in large deviations from the operating schedule and a massive underutilization of Regional Jets. However, previous studies have illustrated that the airline business is competitive and cost minimization alone may not be principal objective behind airline schedule design. Therefore, Formulation II was designed to minimize schedule delay. Schedule delay is inversely proportionally to flight frequency, therefore a minimization of schedule delay translates into a problem of maximizing flight frequency for each segment. Formulation II resulted in a schedule that had a better utilization for smaller aircraft. However, the cost of the schedule generated from Formulation II, was prohibitive. Formulation II, also had a greater sum-of-square error (2.76×10^{10} , about three times larger) than Formulation I due to the fact that the formulation scheduled a very large number of flights on a few segments. The maximum number of flights scheduled is 16,000 which the limit imposed by the formulation. Formulation II also had a drawback of Formulation I, where larger aircraft replaced Regional Jets on many segments. This can be concluded by the fact that most of the points in Figure 5.11 lie under the 45° line. Therefore, Formulation II is not a good representation for the operating flight schedule in the NAS.

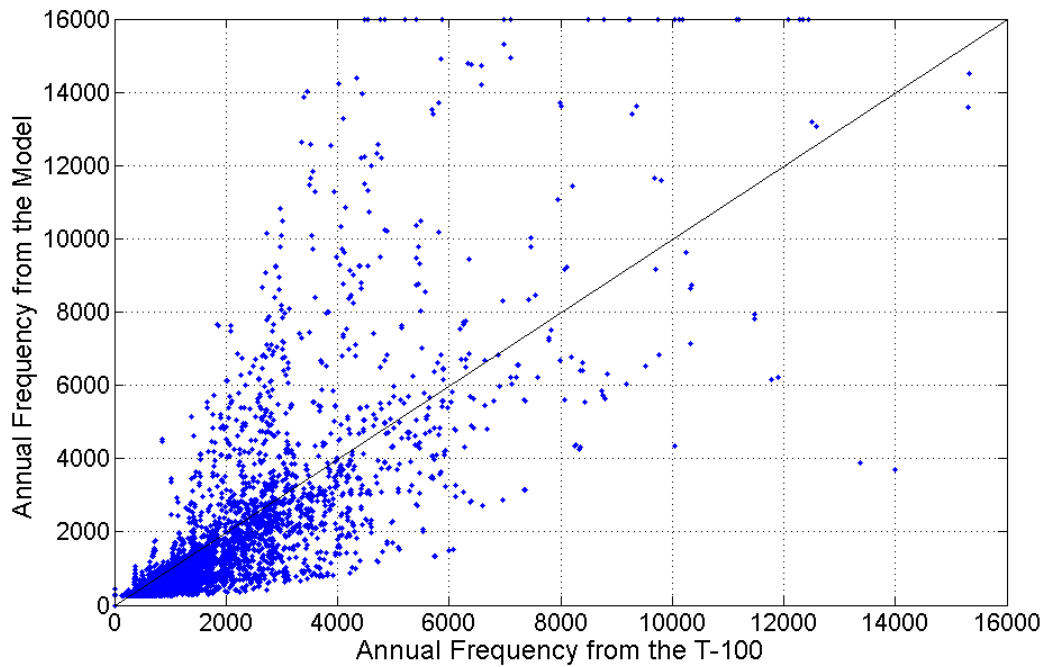


Figure 5.12 Calculated Frequencies from Formulation III vs. Actual Operating Schedule from the T-100.

Formulations I and II, were based on cost minimization and frequency maximization respectively. They suffered from unacceptably low level of service and high service cost. Therefore, Formulation III attempted to achieve a good balance between the two competing objectives. From Figure 5.12, it can be observed that Formulation III has a better performance than Formulations I and II. The sum-of-square error of Formulation III is lower than the error from Formulation II (1.05×10^{10}) and marginally higher than the error from Formulation I and Formulation III also resulted in a more balanced aircraft utilization than Formulation I. Formulation III, however still had an unacceptably low utilization of Regional Jets (by 33%) and a very high frequency on certain segments.

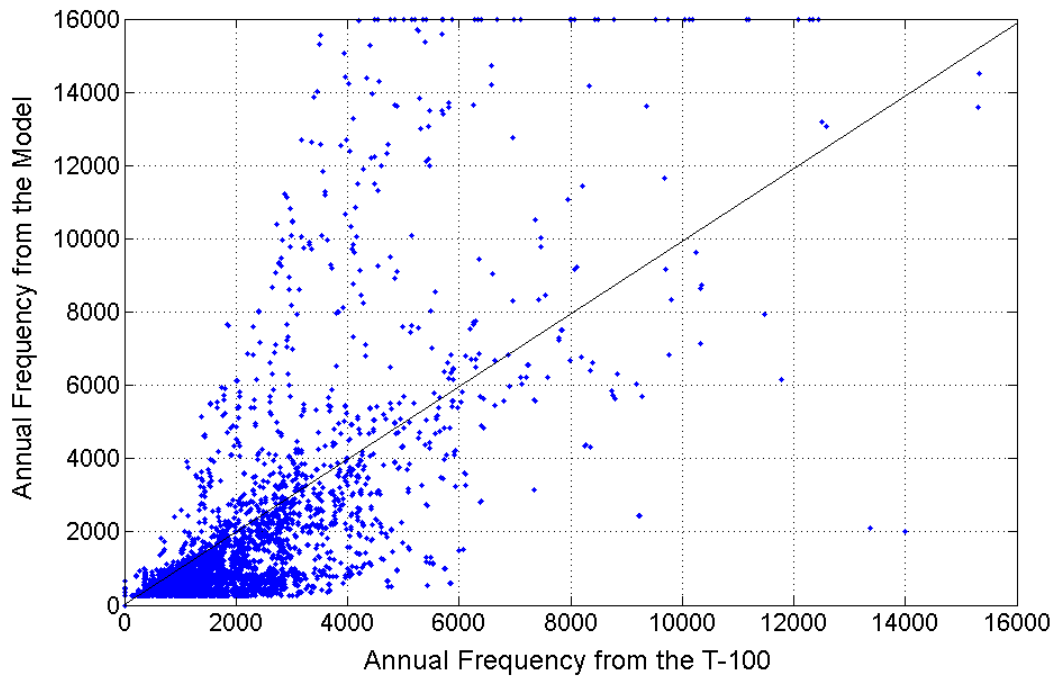


Figure 5.13 Calculated Frequencies from Formulation IV vs. Actual Operating Schedule from the T-100.

Formulation IV used the same approach to flight scheduling as Formulation III. Formulation IV, also attempted to achieve a balance between the objectives of cost minimization and frequency maximization. However, Formulation IV attempted to balance the two objectives only in a subset of the network. Formulation IV was based on the fact that many of the flights from smaller airports are not point-to-point service, but act as a feeder to an airline hub. Formulation IV, was a combination of Formulations I, II and III in which, frequency maximization was the principal objective for segments that served small and non-hubs and a combination of cost minimization and frequency maximization were the objectives for segments that connected large and medium hubs. However, the sum-of-square error was still high (1.47×10^{10}) and some segments still had a very high frequency of service

Formulations I-IV attempted to construct the operating airline schedule without taking into account the existing operating schedule for the baseline time period. This approach is not very desirable as operating schedules can vary significantly between successive time periods. Formulations I-IV also had poor accuracy with respect to the real operating schedule from the T-100.

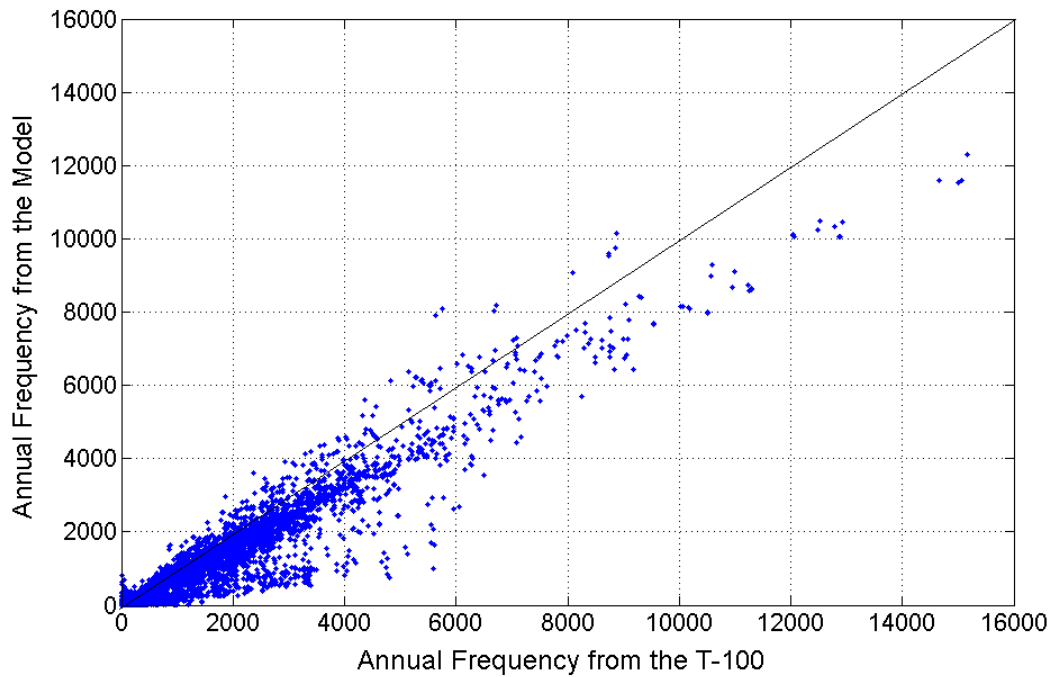


Figure 5.14 Calculated Frequencies from Formulation V vs. Actual Operating Schedule from the T-100.

Formulations I-IV attempted a “clean-slate” approach to the schedule building problem. However, as discussed in the previous section, this approach does not yield good results. Formulations V-VII on the other hand, attempt an incremental approach to the problem of schedule design. The three formulations build the schedule for the next time period by incrementally changing the schedule for the current time period. The main advantage of this approach is that the changes in the operating schedule between two successive time period can be constrained, which is difficult in Formulations I-IV.

Formulation V attempted to minimize the sum of the operating cost and acquisition cost. The formulation also had a set of constraints for limiting the variation in the schedule between successive time periods. From Figure 5.14 it can be observed that, Formulation V performs better than Formulations I-IV. Formulation V also has a lower sum-of-square error (2.61×10^9) than Formulations I-IV. However, Formulations V suffered from the same drawback as formulation I, i.e. an significantly low utilization of Regional Jets (66%).

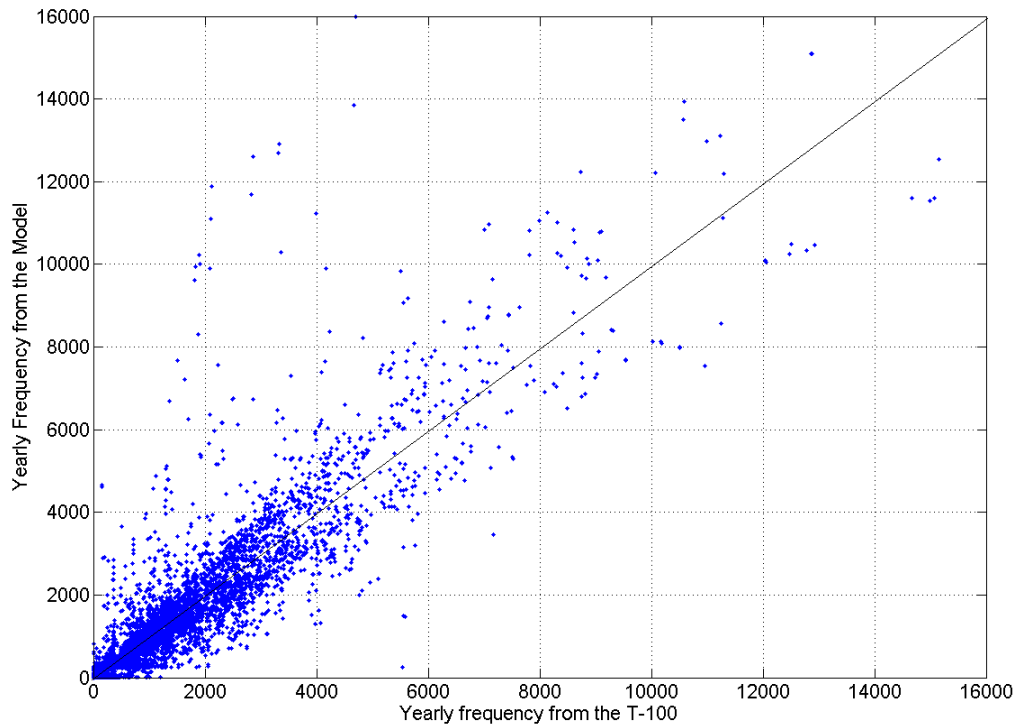


Figure 5.15 Calculated Frequencies from Formulation VI vs. Actual Operating Schedule from the T-100.

Formulation VI attempted to correct the shortcoming of Formulation V, i.e. the underutilization of Regional Jets. However, Formulation VI resulted in a higher sum-of-square error than Formulation V (3.40×10^9). Formulation VI also resulted in very high frequencies for some segments.

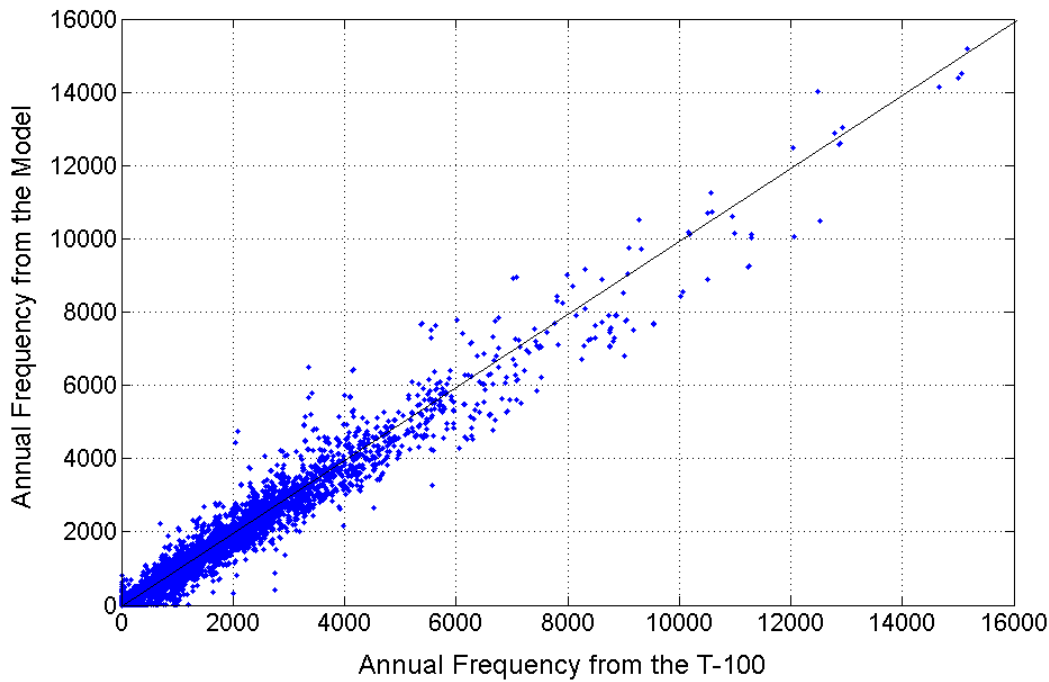


Figure 5.16 Calculated Frequencies from Formulation VII vs. Actual Operating Schedule from the T-100.

Formulation V-VI produced better schedules from Formulations I-IV as measured by their lower sum-of-square error. However, Formulations V and VI still suffered from the same issues as Formulations I and II, i.e. low Regional Jet utilization and a very high frequency of service on some segments. Formulation VII, attempts to minimize the aircraft acquisition costs only. Formulation VII is analogous to finding a basic feasible solution to Formulation I. Formulation VII has the best performance as measured by its sum-of-square error (0.81×10^9) and the degree of utilization of Regional Jets (greater than 99%). Therefore, Formulation VII was adopted.

5.2.2.1) Linear vs. Integer Solution for the Frequency Generator

In Section 4.3.3.2, a Lagrangian based rounding scheme was utilized to derive an acceptable integer solution from the LP relaxation of the original integer program. However, the optimal integer frequencies could be considerably different from the rounding of the LP relaxation. In general, rounding an LP optimal solution is not optimal and is infeasible. The Integer Solution Generator attempts to minimize the degradation in the LP optimal objective function and minimize the infeasibility produced by the rounding scheme. In this section the linear and integer solutions to the problem are compared. A small difference between the linear and integer solutions would indicate that the integer solution generator is an acceptable surrogate, and solving the actual integer program is not necessary. In general, rounding is an acceptable procedure for integer programs with general integer variables, with an order of magnitude in the tens or higher. However, the theoretical validity of LP rounding vs. the integer solution has been debated

in literature, particularly when the input parameters are uncertain. In TSAM, the segment demand used as an input to the frequency generator is uncertain since the demand projections could change considerably depending on the assumptions made. Rappaport (1967) argues that LP based rounding solutions are better than integer solutions particularly under uncertain input parameters. However, Glover (1969) provides an example where none of the rounded LP solutions provide a feasible solution to the original integer program. In this section, the linear and the integer solutions obtained are compared and contrasted

The integer program was solved using CPLEX 11.0. The original formulation was very difficult to solve in CPLEX as the formulation had a number of tightly bound variables and variables that were fixed at zero. Formulation VII therefore was re-formulated by removing the fixed and tightly bound variables from the formulation. The new formulation had about 28,000 variables instead of 47,000. Out of these 28,000 variables about 12,000 variables had zero values in the LP relaxation and another 9,000 variables had integer values. The number of fractional variables from the LP relaxation was about 7,000. The objective of the formulation was to find an integer feasible solution with a reasonably small LP-IP gap, as the problem had multiple integer solutions. The final integer solution chosen had a 4.61% difference with the optimal LP solution. The integer solution was obtained in about 4.5 hours on a dual core Intel Centrino machine. The comparisons of the linear and integer solutions are plotted as a two-tuple $(N_{ij}^*$ and N_{ij}^{**} , where N^{**} is an optimal integer solution), in Figures 5.17 – 5.26.

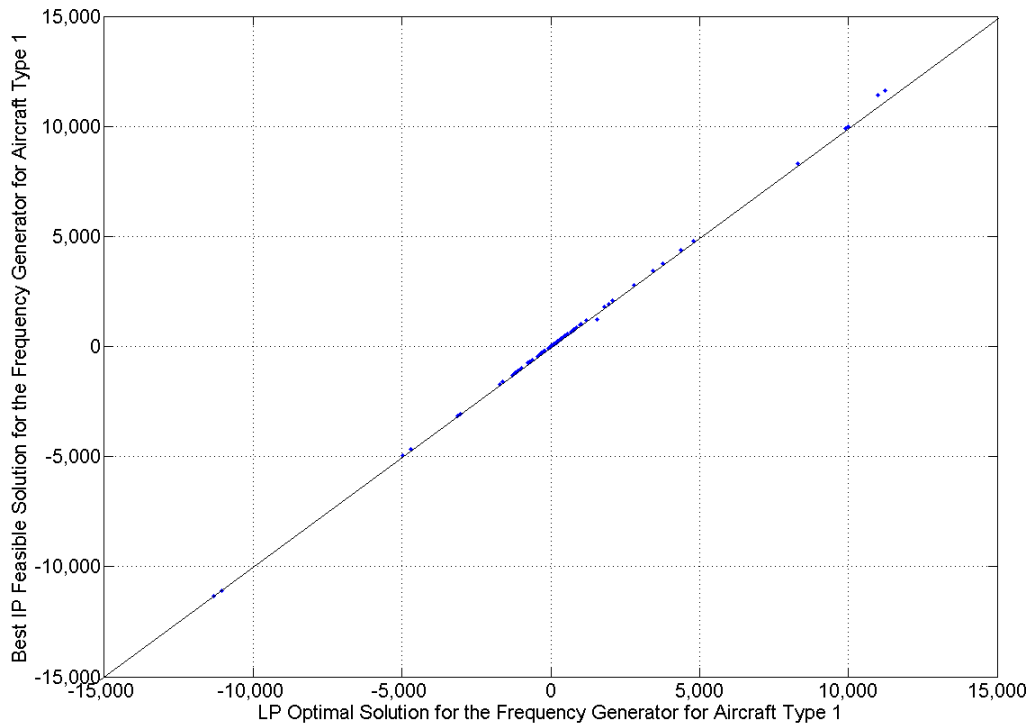


Figure 5.17 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 1.

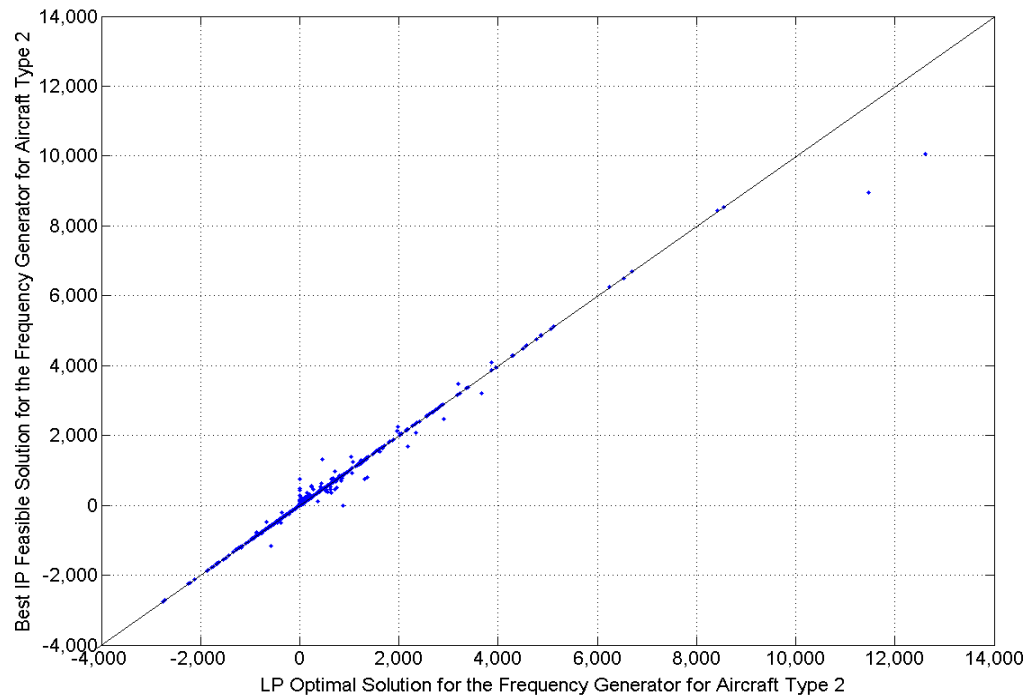


Figure 5.18 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 2.

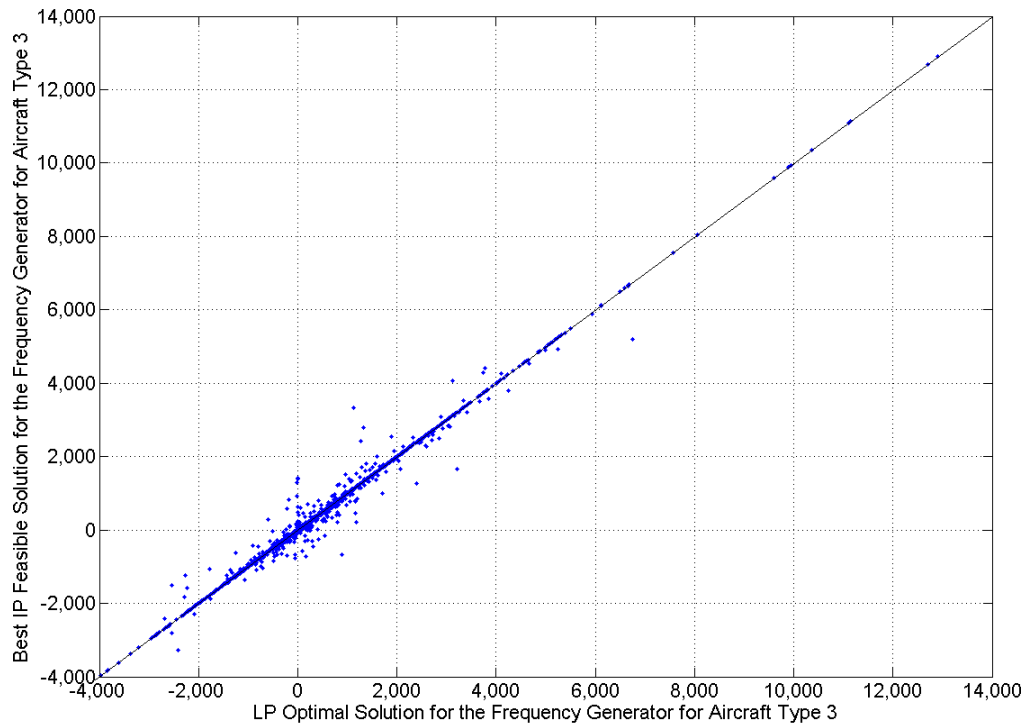


Figure 5.19 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 3.

It can be observed from Figure 3, that the difference between the optimal linear solution and an integer feasible solution is higher for aircraft of type 3, than for other aircraft groups. This can be attributed that aircraft group 3 consists of Regional Jets, which typically have a higher utilization than aircrafts of other groups. Therefore, it could be argued that the linear – integer solution gap is higher due to the larger solution space for aircraft group 3. The same phenomenon is observed for aircraft group 4.

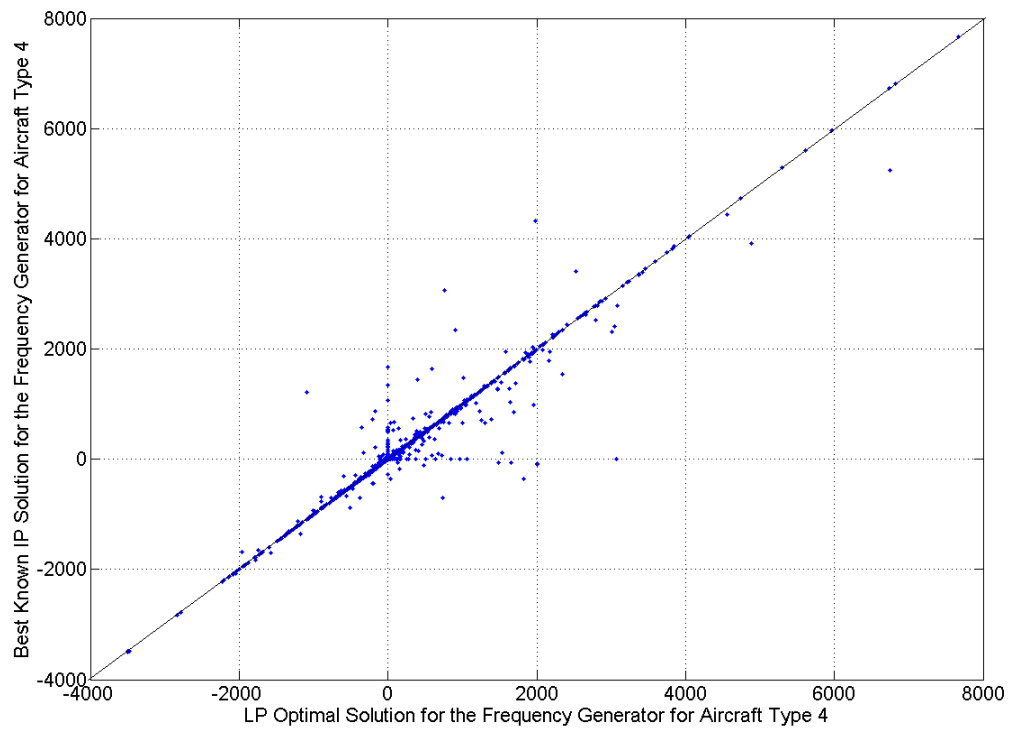


Figure 5.20 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 4.

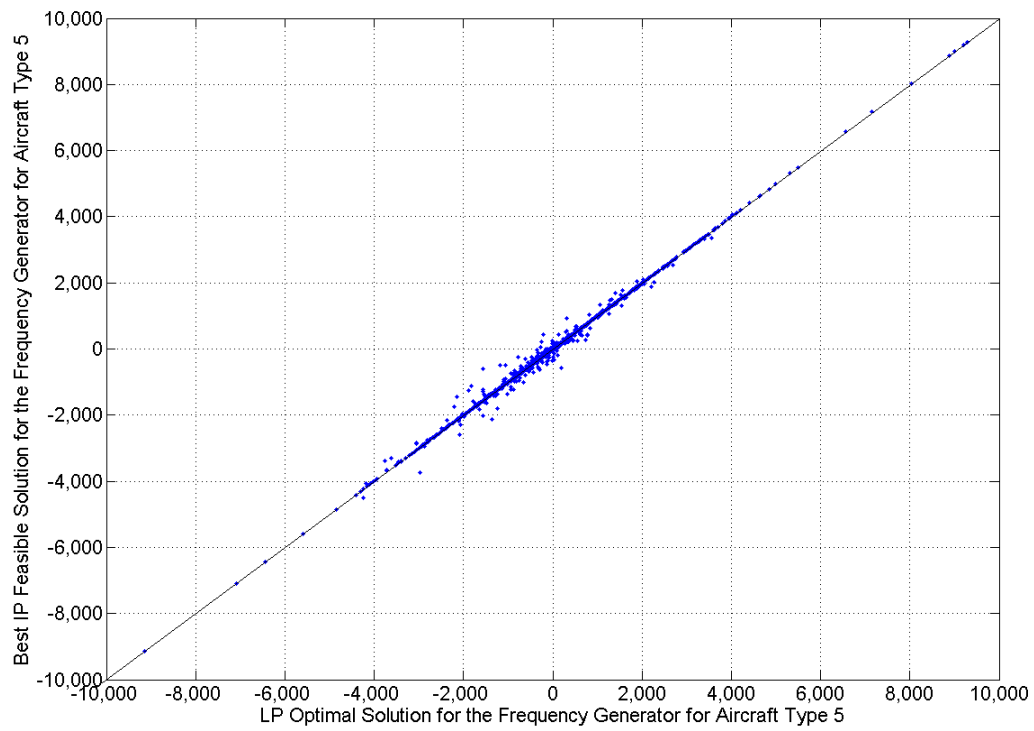


Figure 5.21 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 5.

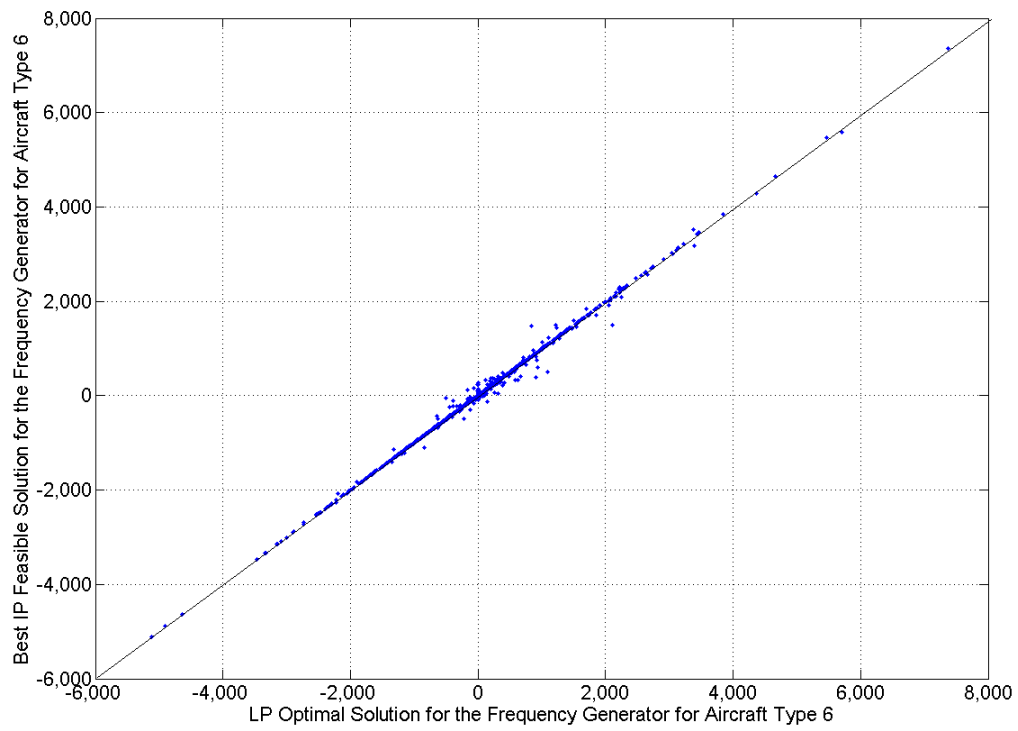


Figure 5.22 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 6.

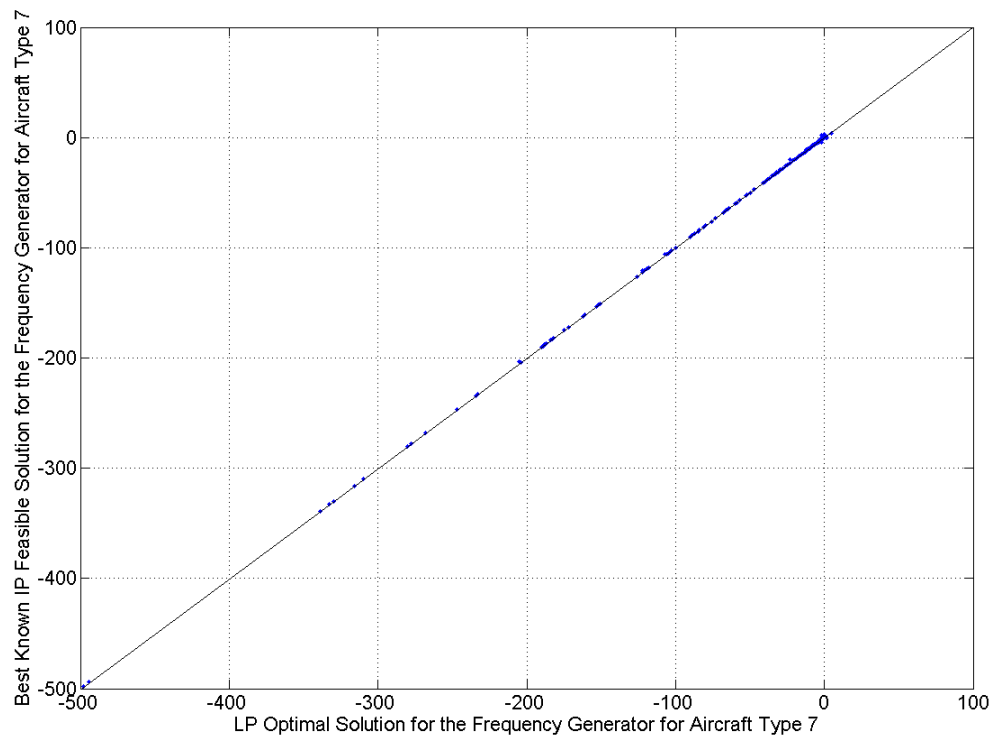


Figure 5.23 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 7.

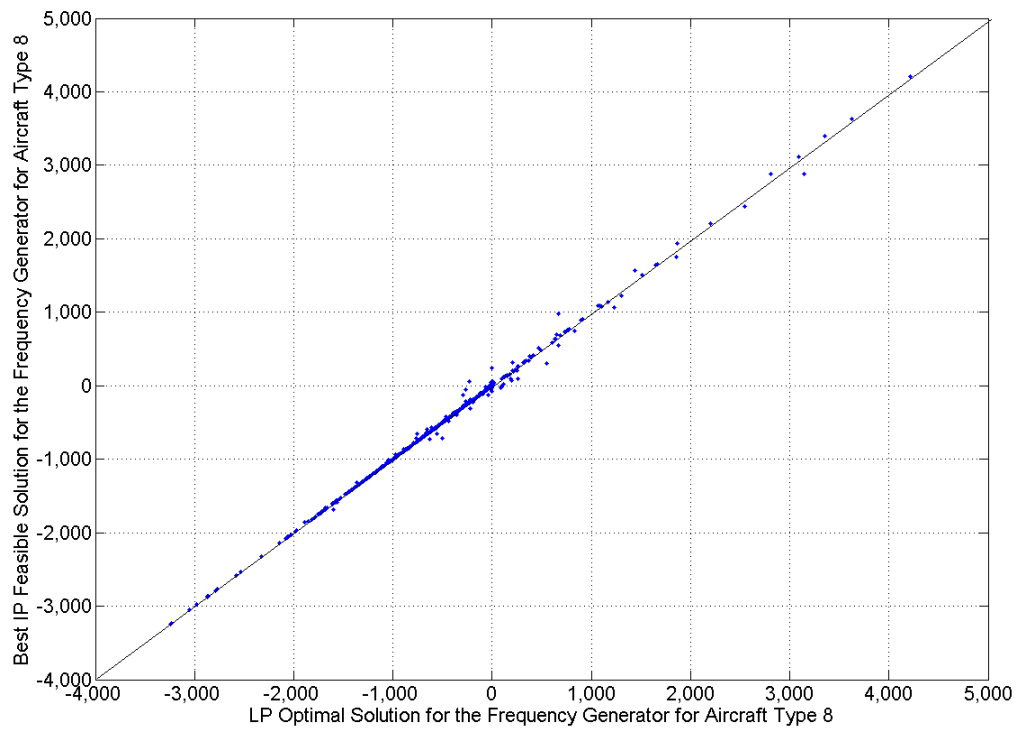


Figure 5.24 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 8.

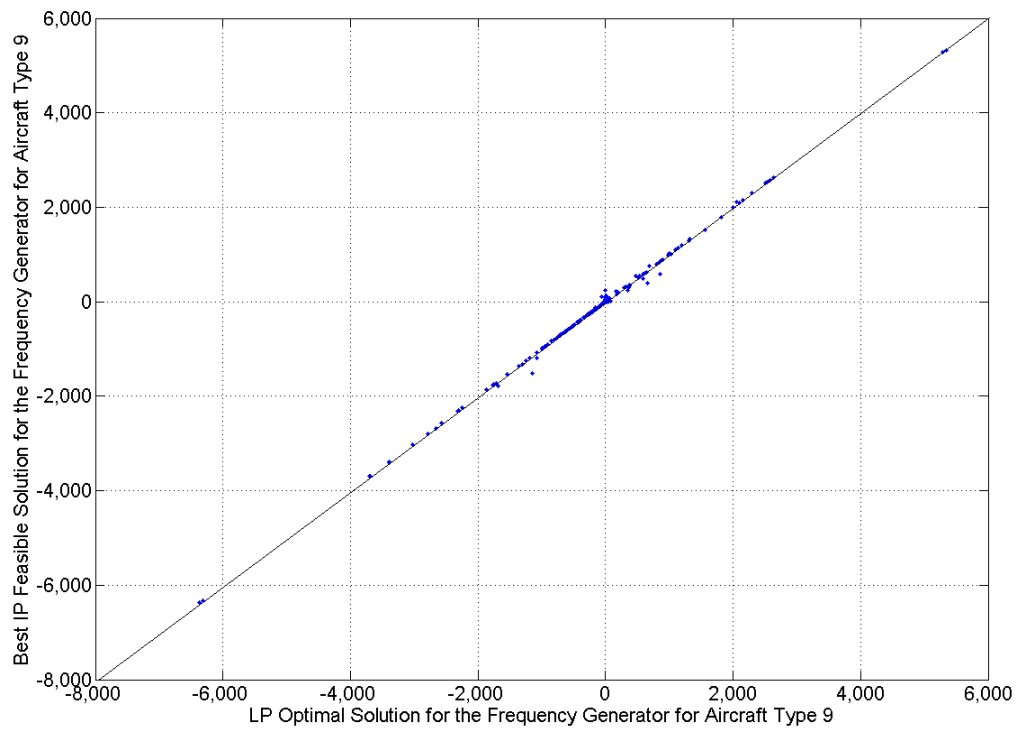


Figure 5.25 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 9.

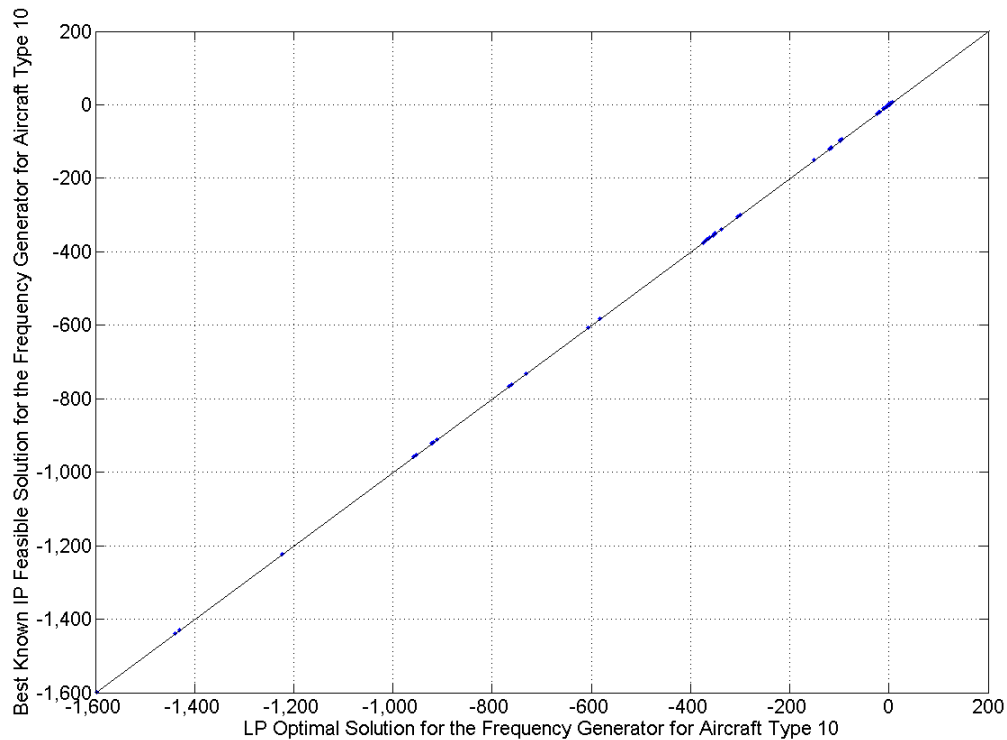


Figure 5.26 The LP Optimal Solution vs. Integer Solution for the Frequency Generator for Aircraft Type 10.

From the Figures 5.17 – 5.26 it can be seen that the linear programming optimal is close to the best known integer feasible solution. The error between the integer and linear solutions is very small for aircraft type 1 and large for aircraft type 4. Since the objective of the problem was frequency maximization and cost minimization, it could be argued that optimal solution would tend to underutilize larger aircraft. It can also be concluded that the error between the linear and integer solutions is larger for aircraft with a higher utilization rate. For example, the integer solution is practically the same as the rounded LP relaxation solution for aircraft types 1,7 and 10 which have low utilization rates. On the other hand the errors are larger for aircraft types 3, 4 and 5 which have higher utilization rates. However it can be concluded that, the overall LP rounded solution is a good surrogate for the actual integer solution and the integer solution generator can be used to obtain good integer estimates for the frequency generator.

5.2.3) The Network Evolver

As mentioned in Section 4.4, the network is evolver is composed of five steps.

- (1) Cutting service on non-profitable routes
- (2) Adding new point-to-point service
- (3) Adding new hub-to-spoke connections

- (4) Adding and removing hubs
- (5) Introducing service from airports that have no significant commercial service at present

The sub-module to cut service on unprofitable routes utilizes a binary probit regression scheme to derive the likelihood of cessation of service. The probability of the cessation of service is then converted into a binary decision variable through the use of a threshold value. Since the output of the model is a binary variable, common statistical tests for accuracy, for example the R^2 , cannot be used as they are defined only for continuous response variables. The binary response variable can take a value of zero or one depending on whether the segment is kept in service or deleted from the network.

The model accuracy is tested using a “hit-rate” statistic. The hit rate is derived by computing the ratio of the number of times the binary output from the model matches the binary response from the data. Higher hit-rates imply higher model accuracy. The hit rate also is dependent on the threshold value that is utilized to convert the continuous probability value to a binary response. The threshold values were fixed at values that maximized the hit rates. Figure 5.27 illustrates the variation of the hit-rate with the threshold value for the model that cuts unprofitable routes. Another objective of the model is to match the number of routes cut by the model to the number of routes cut by the airlines. The results of the probit model are illustrated in Table 5.6.

Table 5.6 Coefficients and Hit Rates for the Binary Probit Module to Cut Unprofitable Routes.

Segment Type	Variable						
	Service Frequency	Profit Per Seat	Load Factor	Log Of Distance	Essential Air Service	Threshold	Hit Rate
LH-MH/ MH-LH	-0.0008346	-0.005115	-1.8659	N/A	N/A	0.13	0.9670
LH-SH/ NH-SH /NH -LH	-0.0008735	-0.003003	-3.7644	0.22062	-0.6784	0.24	0.9270
MH-MH	-0.0004143	N/A	-1.7665	N/A	N/A	0.20	0.9535
MH- NH /SH-NH /SH - MH	N/A	-0.015201	-3.1252	N/A	-2.30206	0.25	0.8864
NH/SH – SH/NH	-0.001111	N/A	-2.8521	0.2658	-0.6259	0.40	0.7513

The second part of the network evolver module attempted to add segments to the network based on the passenger demand. There are two types of links that could be added to the network, point-to-point links that serve primarily O-D passengers and hub-spoke links whose principal traffic is from connecting passengers. The probability of adding a new

hub-spoke link is computed by a binary probit regression model. The probabilistic output of the model is then converted to a binary response by the use of a threshold value. The results of the model are illustrated in Table 5.7.

Three probit models for the addition of new hub-spoke connections were calibrated. The first model utilized the total passenger demand from the spoke, hub-spoke distance, maximum frequency from the spoke, number of spokes connected to the hub and a binary variable for the hub-connectivity of the spoke airport as explanatory variables. The second model removed the number of spokes as an explanatory variable. The third model used a slightly different set of airline hubs from the first model.

Table 5.7 Coefficients and Hit Rates for the Binary Probit Module for the Introduction of New Hub-Spoke Connections.

Model	Variable						
	Hub-Spoke Distance	Existing Hub Connection	Max. Spoke Frequency	No.Of Spokes	Total Demand	Threshold	Hit Rate
Model 1	0.003234	1.8796	-0.0003298	-0.0659	-7.302×10^{-7}	0.60	0.992
Model 2	0.003234	1.7135	-0.0003350	N/A	-1.311×10^{-6}	0.60	0.994
Model 3	0.003242	1.7414	-0.0003381	N/A	-1.351×10^{-6}	0.60	0.992

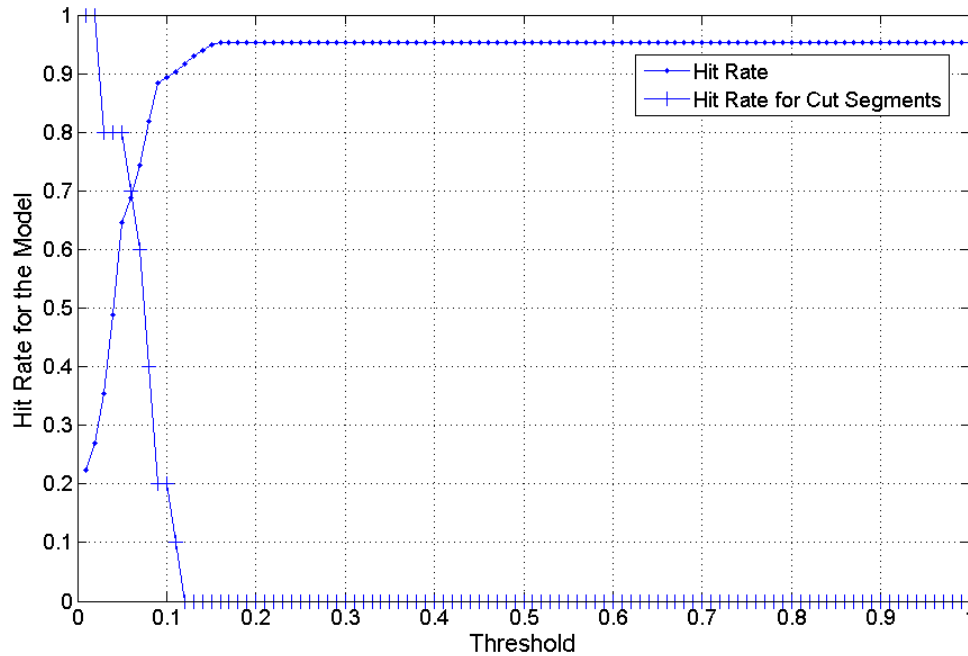


Figure 5.27 Variation of the Total Hit Rate and the Hit Rate for Cut Segments with the Threshold Value.

5.2.4) The Fare Prediction Module

As mentioned in Section 4.5, the fare prediction module derives the fares for each route in the network, since the fare is an important exogenous variable to predict route flows. The fare prediction problem can be subdivided into two distinct sub-problems; the first sub-module derives fares for the routes for which the DB1B has insufficient information and the second sub-module projects the baseline year fares into the future.

The fare sub-model would use a pure distance based model which can be expressed as,

$$Fare_{ijk} = \frac{1}{(a + b * D_{ijk}^c)} \quad (344)$$

Where

a, b, c : Constants

D_{ijk} : Flown distance between airports i and j through airport k .

$Fare_{ijk}$: Fare between airports i and j through airport k .

Separate fare models were calibrated for separate hub types.

The calibration results for large and medium hubs are illustrated in Figures 5.28 and 5.29

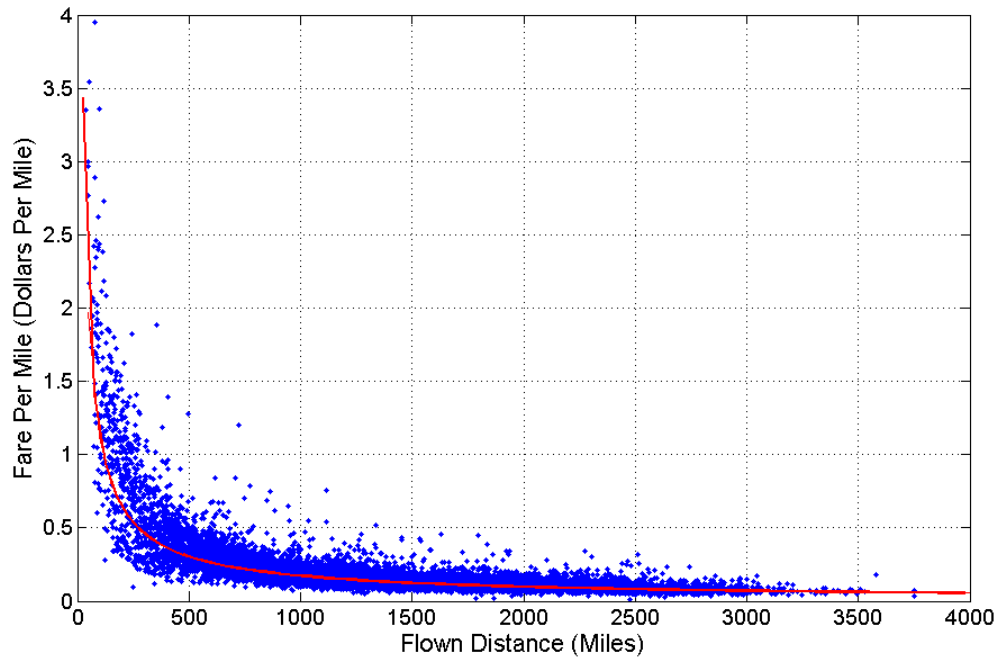


Figure 5.28 Variation of Fare per Mile and Distance for Coach Class for Passengers Originating in the Large Hubs.

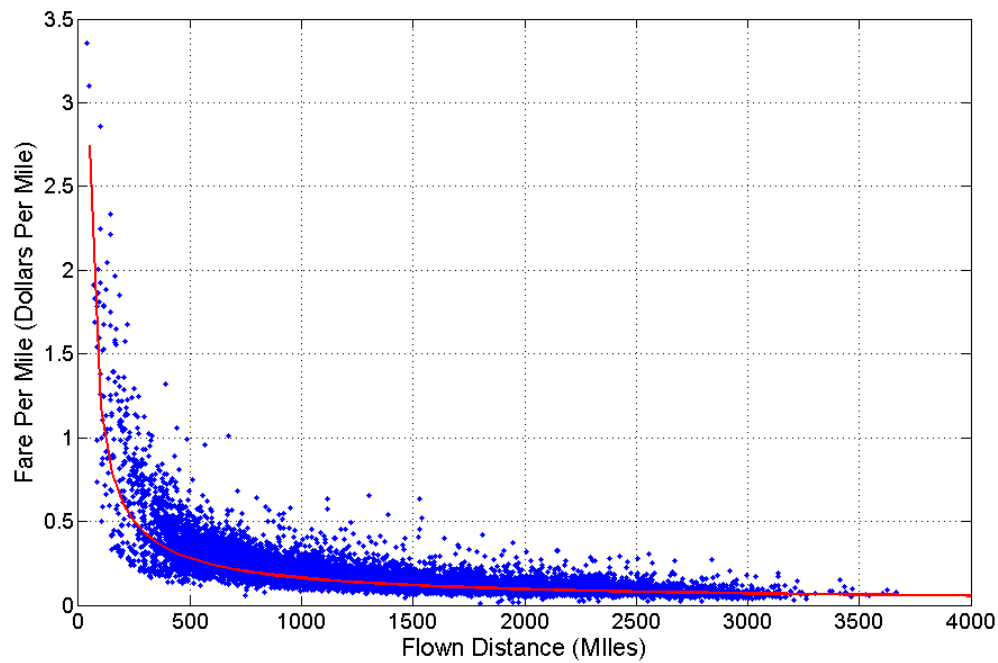


Figure 5.29 Variation of Fare per Mile and Distance for Coach Class for Passengers Originating in the Medium Hubs.

The second part of the fare prediction module attempts to predict fare trends into the future. The trends are predicted by a logarithmic regression over a few socioeconomic and airline related explanatory variables. The variables used in the model are

1. Demand
2. Load Factor
3. Dummy for 9/11 terror attacks
4. Cost per seat
5. Distance

Distance was implicitly included by performing a separate trend-line analysis for each distance bracket. The results of the trend-line analysis are shown in Figures 5.30 and 5.31 and the regression coefficients are shown in Table 5.8

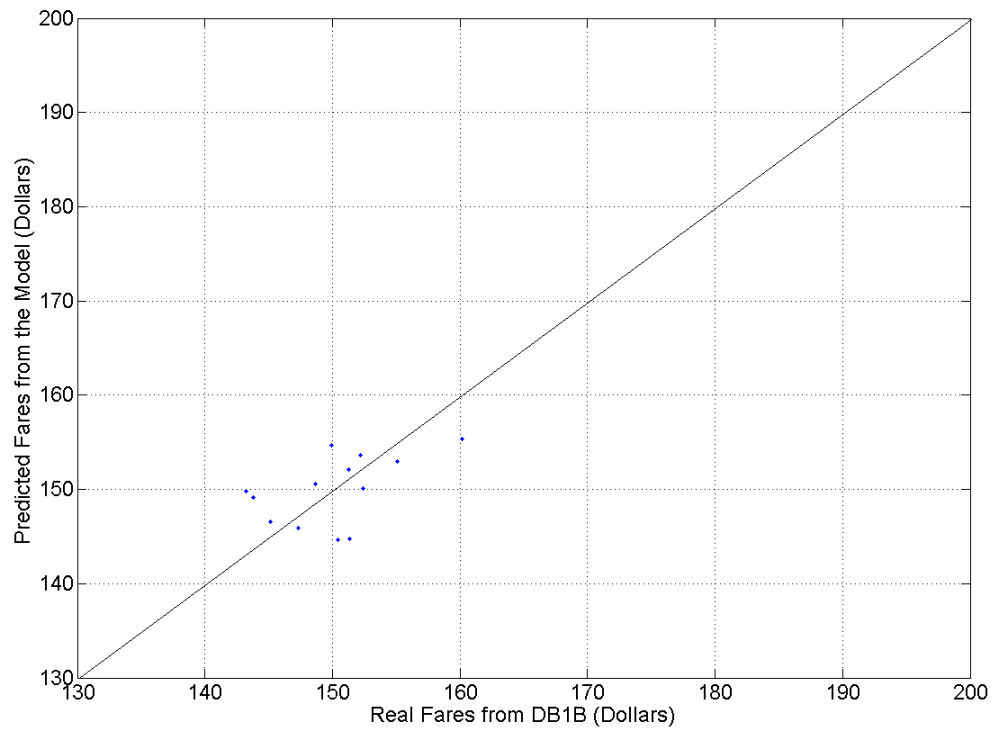


Figure 5.30 Average Real vs. Average Predicted Fares for all O-D Pairs with a Great Circle Distance less than 500 Miles for the Years 1993-2006.

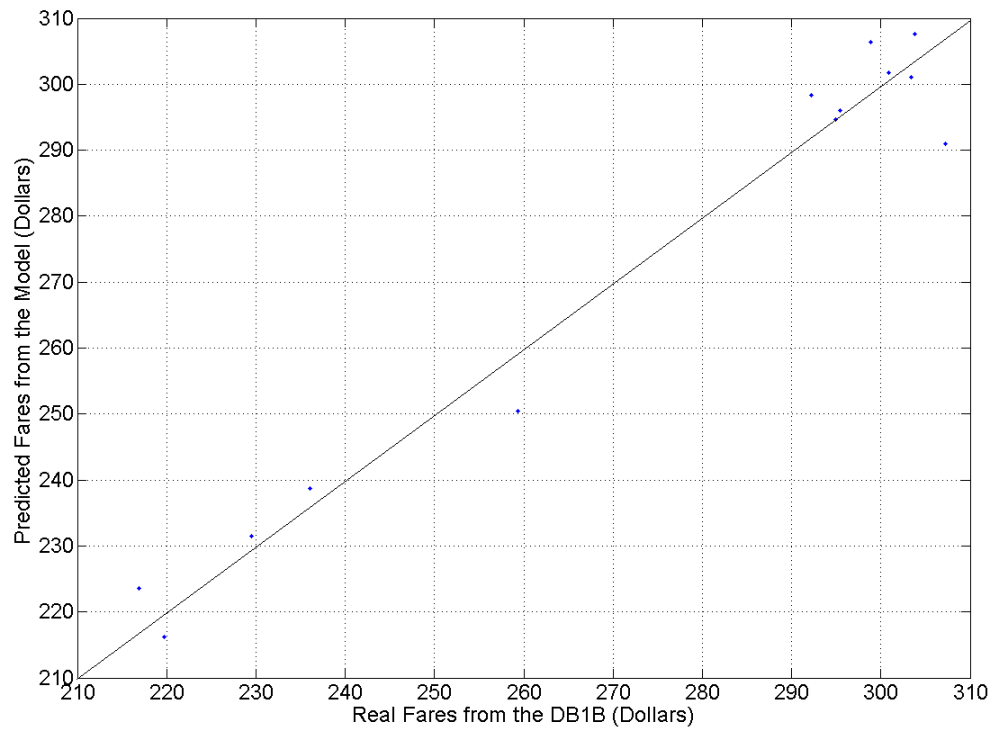


Figure 5.31 Average Real vs. Average Predicted Fares for all O-D Pairs with a Great Circle Distance more than 2500 Miles for the Years 1993-2006.

Table 5.8 Linear Regression Coefficients for the Fare Trend Model.

Distance Bracket (miles)	Coefficients		
	Intercept	Load Factor	9/11 Dummy
0-500	103.38	21.49	-4.39
500-1000	277.80	-155.88	-18.97
1000-1500	315.62	-182.03	-22.23
1500-2000	375.78	-227.29	-32.98
2000-2500	393.41	-241.00	-43.76
2500-3000	440.74	-268.09	-54.30

5.3) Integrated Model Results

In this section, the results from the integrated model run are presented. The supply side model of the NAS as described in Section 4, is run concurrently with the demand side model in TSAM. However, the demand side model in TSAM could not adequately capture the NAS segment demand. As the results of the integrated runs are predicated on the demand side model of TSAM, the integrated model was run with the T-100 segment demand as the baseline and the demand was grown using the TSAM demand growth rate. The integrated model was run from 2006 to 2010.

5.3.1) Mode Choice Trends

In this section, the trends in total travel demand in the continental United States are presented. Tables 5.9 and 5.10 illustrate the total air and auto trips from 2006 to 2010.

Table 5.9 Total CA Trips (One-Way).

Year	One-Way CA Trips
2006	445,747,518
2007	397,061,412
2008	401,956,313
2009	401,936,068
2010	416,650,974

Table 5.10 Total Auto Trips (One-Way).

Year	One-Way Auto trips
2006	967,857,444
2007	1,012,959,795
2008	1,031,388,439
2009	1,052,268,884
2010	1,066,009,450

The demand drops significantly from 2006 to 2007 due to the airlines cutting a large number of unprofitable routes. After 2007, the demand once again follows an increasing trend due to the growth in population and economy.

5.3.2) Origin-Transfer-Destination (O-T-D) Trends

The mode choice demand matrices obtained for every year are used to derive the segment flow and enplanements for every airport in the network. The enplanements follow the

same trends as the mode choice demand, indicating that the average number of legs per O-D pair is stable. From Table 5.13 it can be inferred that the average number of legs per O-D pair remains constant over the simulation period.

Table 5.11 Total CA Domestic Enplanements (TSAM).

Year	Total CA Enplanements (Domestic)
2006	629,785,343
2007	587,200,736
2008	593,609,072
2009	615,941,218
2010	666,046,921

Table 5.12 Average Number of Legs Per O-D Pair.

Year	Average CA Legs (Domestic)
2006	1.410
2007	1.478
2008	1.478
2009	1.532
2010	1.598

Table 5.13 Proportion Of Direct Passengers.

Year	Proportion of Direct Passengers
2006	73.38%
2007	72.56%
2008	70.87%
2009	71.46%
2010	67.51%

5.3.3) Network Evolver Trends

The network evolver trends are derived from the segment and O-D demand in TSAM. Since TSAM is executed in a phase lagged fashion, the schedule for the year 2006 is the base template for the year 2006 and 2007. For the year 2006 and 2007, segment costs are computed using the original costs for each individual aircraft. For the years 2008 and 2009 costs are computed from the schedule produced by the schedule generator.

Table 5.14 Network Revenue, Costs and Profit.

Year	CA Network Revenue (Billion \$)	CA Network Cost (Billion \$)	CA Network Profit (Billion \$)
2006	84.29	99.74	-15.45
2007	78.26	96.91	-18.65
2008	82.59	106.97	-24.37
2009	82.57	98.34	-15.77
2010	83.23	100.59	-17.36

From table 5.15 it can be seen that the airlines do not break even in the domestic market. This can be attributed to soaring fuel costs from the year 2006 to 2008 and the consequent slackening demand. From the Air Transport Association (ATA), it is observed that domestic carriers made a net profit in the years 2006 and 2007. However, the domestic carriers revenues include both domestic and international operations, and international operations are typically more profitable. Therefore, for a consistent comparison it was necessary to compute the total revenue and cost for the domestic segment alone. From the ATA data on passenger yields the total domestic revenue for the years 2006 and 2007 was determined to be \$76.2 billion and \$ 76.8 billion dollars respectively (2006,2007) which agree well with the model predicted values for the same years. From the form P-6 the total domestic operating cost was \$ 97 billion and \$ 101 billion which again are approximately the same as the model predicted values.

Table 5.15 Number of Segments in the CA Network.

Year	Number of Segments in the CA Network	Segments Cut by the Network Evolver	Segments Added by the Network Evolver
2006	4972	367	440
2007	4429	29	242
2008	4382	90	127
2009	4296	32	76
2010	4182	26	103

From Table 5.16 it can be inferred that the model ceases service in the unprofitable segments till almost all the segments are profitable or otherwise attractive to operate. Therefore, the model culls a larger number of segments in the first year and then the number of cut segments progressively decreases to a more stable number.

Table 5.16 List of Airline Hubs by Year.

Year	Number of Airline Hubs
2006	28
2007	27
2008	25
2009	25
2010	26
2011	24

Table 5.17 lists the number of airline hubs for each year. In this model an airline hub is defined as any airport that has direct flights to at least 53 other airports and 15 small and non-hubs. Table 5.18 contains the specific airline hub ID by year.

Table 5.17 List of Airline Hubs by FAA Three Letter Code.

Airline Hubs (2006)	Airline Hubs (2007)	Airline Hubs (2008)	Airline Hubs (2009)	Airline Hubs (2010)
ATL	ATL	ATL	ATL	ATL
BOS	BOS	BOS	BOS	BOS
CLT	CLT	CLT	CLT	CLT
CVG	CVG	CVG	CVG	CVG
DCA	DCA	DCA	DCA	DCA
DEN	DEN	DEN	DEN	DEN
DFW	DFW	DFW	DFW	DFW
DTW	DTW	DTW	DTW	DTW
EWR	EWR	EWR	EWR	EWR
IAD	IAD	IAD	IAD	IAD
IAH	IAH	IAH	IAH	IAH
LAS	LAS	LAS	LAS	LAS
LAX	LAX	LAX	LAX	LAX
LGA	LGA	LGA	LGA	LGA
MCO	MCO	MCO	MCO	MCO
MSP	MSP	MSP	MSP	MSP
ORD	ORD	ORD	ORD	ORD
PHL	PHL	PHL	PHL	PHL
PHX	PHX	PHX	PHX	PHX
PIT	PIT	PIT		
SEA	SEA	SEA	SEA	SEA
SFO	SFO			
SLC	SLC	SLC	SLC	SLC
STL	STL	STL	STL	STL
CLE	CLE	CLE	CLE	CLE
MDW	MDW			
MEM	MEM	MEM	MEM	MEM
MKE			TPA	TPA
				FLL

From Table 5.15, some of the prominent airports that lose their status as an airline hub are Pittsburgh International (PIT), Chicago Midway (MDW), and Milwaukee General Mitchell Airport (MKE). Pittsburgh International lost its status as an airline hub following the decision of US Airways to downgrade the airport to a focus city in 2004. The number of direct daily flights from PIT dropped from 77 to 46 and most of the smaller destinations lost direct service from Pittsburgh (Wikipedia,2008). Chicago Midway lost its status as an airline hub following ATA Airlines decision to withdraw from Midway (Wikipedia,2008) and cease all operations.

From Table 5.15, it can be observed that Tampa International (TPA) attains the status of an airline hub in 2009 and Fort Lauderdale –Hollywood International (FLL) attains the status of an airline hub in 2010. This is corroborated by the fact that Spirit Airlines made FLL a hub in 2002, however Spirit Airline's presence in FLL received a significant boost in 2007, making Spirit Airlines the largest carrier in FLL. Therefore, it could be argued that FLL attained the status of an airline hub in 2007.

Another issue which is under debate is the fate of the Essential Air Service Program. The federal government has plans to slash the budget to \$50 million, less than half the present funding (Wikipedia, 2008). The network evolver module has a dummy variable to account for the effects of the essential air service subsidy. The model does not assume that the essential air service subsidy would be permanent and accounts for the fact that airlines would cut service to many small airports when the losses from the service to the airport exceed the government subsidy. Most of the airports that lose service from 2006 to 2009 fall in the essential air service category.

Table 5.18 Partial List of Airports Disconnected from the Network.

Airport ID	Hub Type
ABR	4
AIA	4
ALS	4
ART	4
BFD	4
BHB	4
BKX	4
BLF	4
BRD	4
BRL	4
CDR	4
CEZ	4
CGI	4
CIC	4
CKB	4
CNM	4
CNY	4
CLD	4
CVN	4
DIK	4
DVL	4
EAR	4
EAT	4
ELD	4
ESC	4
EWB	4
FHU	4
FKL	4
FOD	4
GBD	4
GCK	4
GDV	4
GGW	4
GLH	4
GRI	4
HIB	4
HII	4

The changes in network topology is depicted in from Figures 5.32 to Figures 5.39

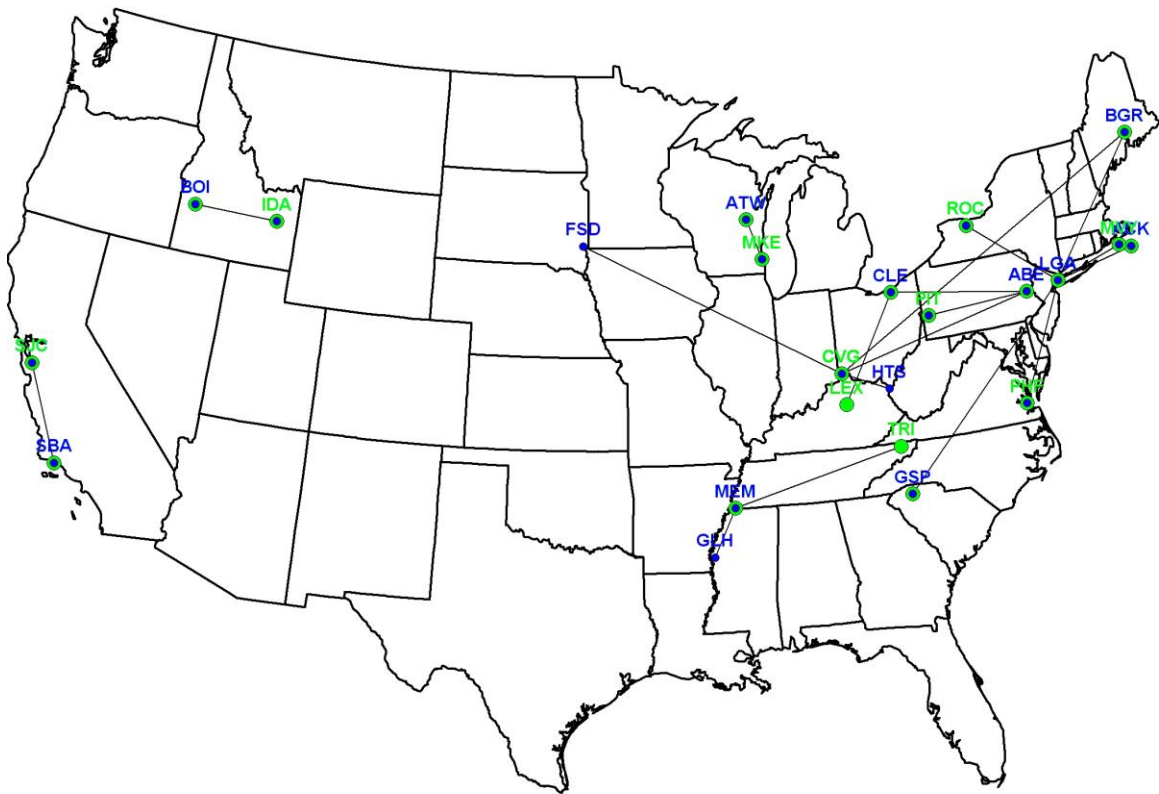


Figure 5.32 Map of Segments Cut in 2007.

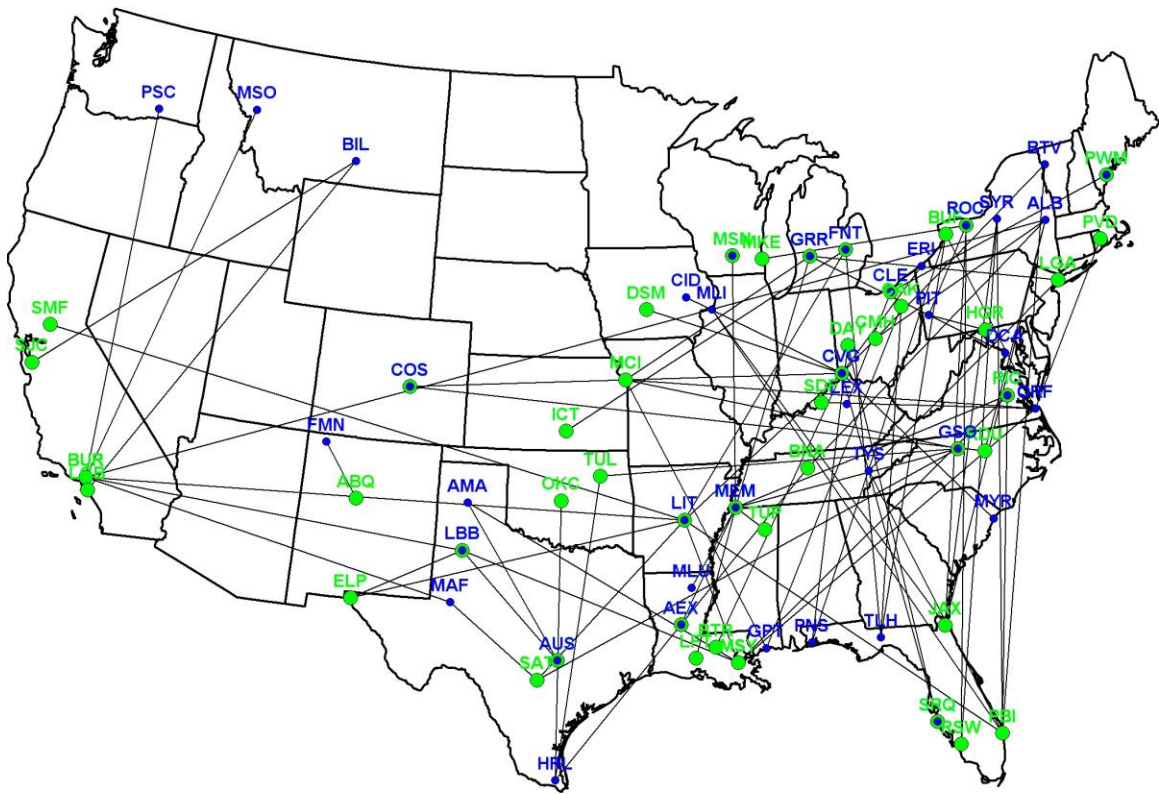


Figure 5.34 Map of Segments Cut in 2008.

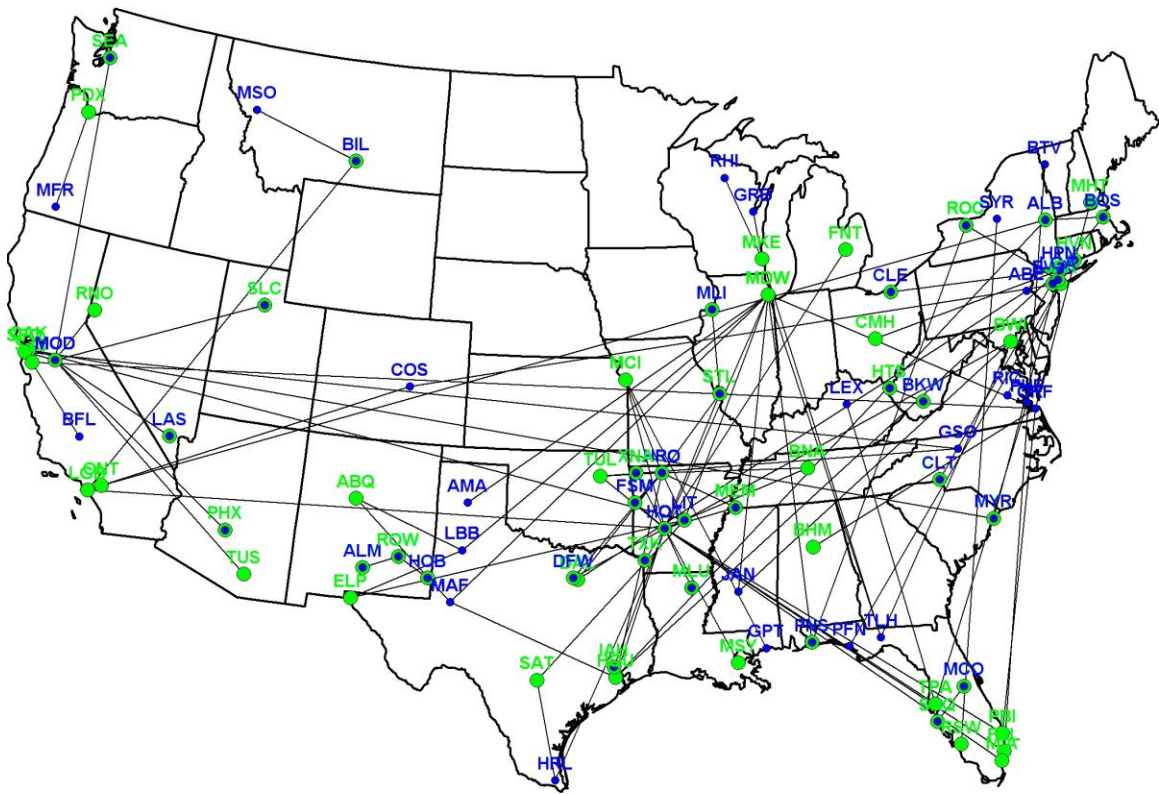


Figure 5.35 Map of Segments Potentially Added in 2008.

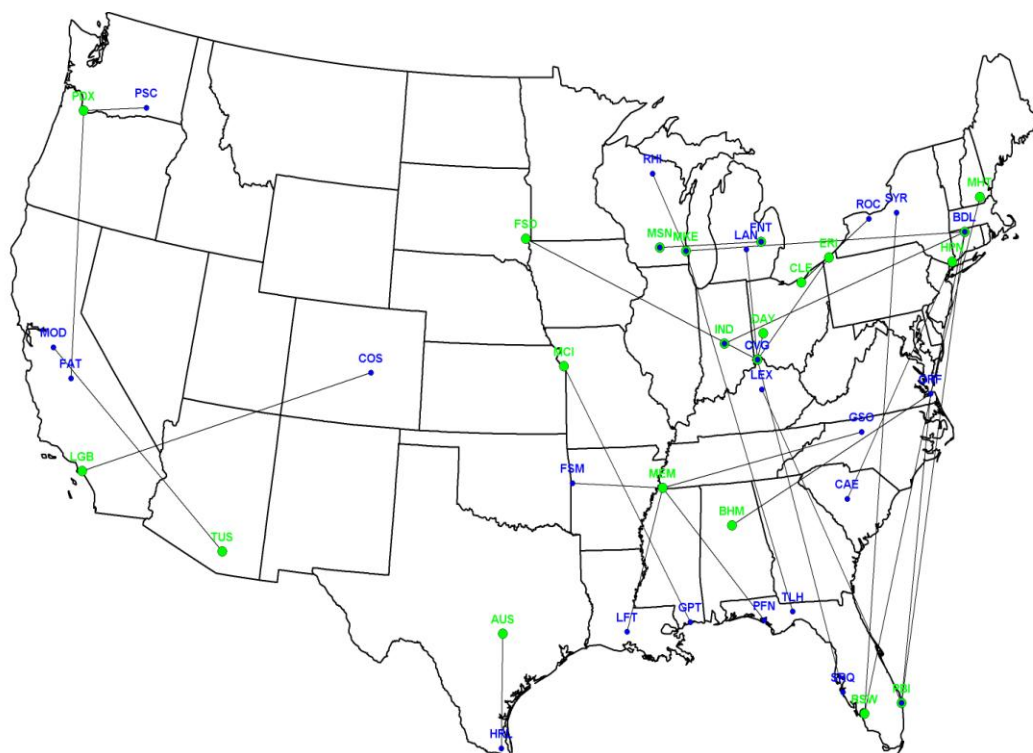


Figure 5.36 Map of Segments Cut in 2009.

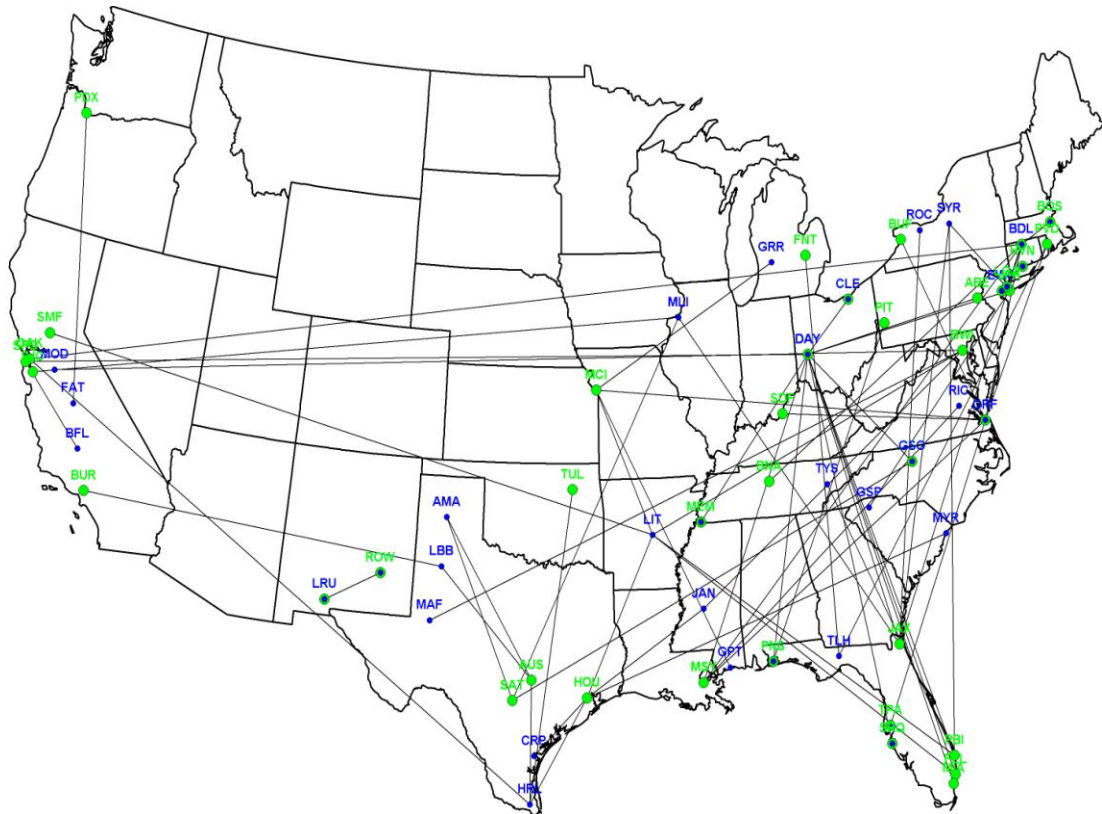


Figure 5.37 Map of Segments Potentially Added in 2009.

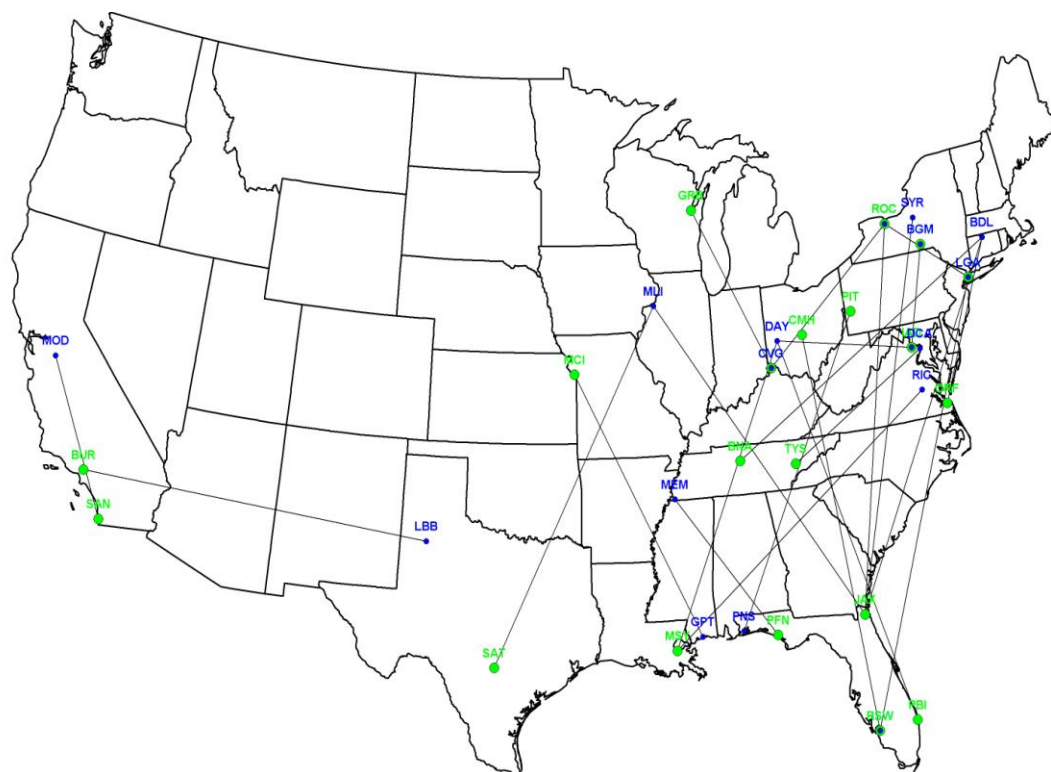


Figure 5.38 Map of Segments Cut in 2010.

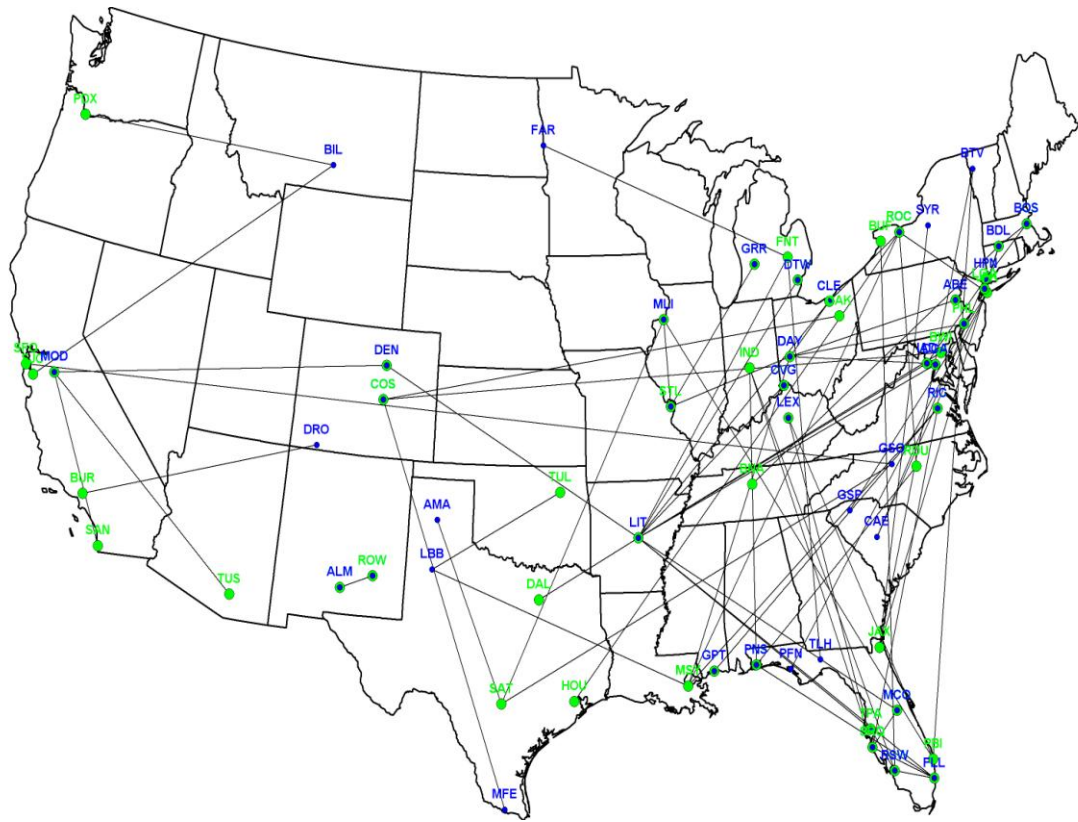


Figure 5.39 Map of Segments Potentially Added in 2010.

Chapter 6 Conclusions and Recommendations

In this dissertation, an attempt has been made to develop a scheme to reliably predict the trends in the airline industry given the trends in airline passenger demand. The dissertation tackles the problem of predicting the airline industry trends into four sub-problems and solves each of the four sub-modules separately. The four sub-modules are finally integrated into the Transportation System Analysis Model (TSAM). The dissertation completes the supply side modeling of TSAM and therefore, enables an analyst to obtain the complete view of the National Airspace System (NAS). To the knowledge of the author none of the models in literature, except Niedringhaus (2004) attempt to model the entire NAS with supply and demand interactions. The model acting in concert with TSAM can also facilitate evaluation of new technologies in aviation and predict the emergence of new airline hubs, congestion at major airports and trends in airline fleet size and type. In the last chapter, several techniques to tackle each of the sub-modules have been presented. In this chapter, the final technique chosen and the justification for choosing the technique are presented along with recommendations for future research.

The Air Traffic Network Flow Model (O-T-D Model) uses Travel Time, Fares, Frequency, Seats and Number of Leg for 443 airports and a network based on the OAG to construct a multinomial logit model. The best model is the multinomial logit with the airport delay variable as the model was sensitive to the capacity constraint at the connecting airport. The nested logit model did not have a significantly better log-likelihood ratio and the inclusive values for the nested logit are close to unity. Therefore, it is concluded that a nested logit would not yield any appreciable increase in the predictive power of the model. The model derives more positive results (Craig-Uhler $R^2 = 0.9$) for segment flows between large and medium hubs, which constitute about 75% of the segment flows in the NAS (2005). The model does not yield very good results (Likelihood Ratio = 0.4) from small and non-hubs. However, the DB1B is also not a good source of information for the smaller airports and this could be one of the reasons for the model's performance for the small hubs. The model gives fairly accurate estimate of segment flows and airport enplanements for the large and medium hubs. The inclusion of the airport delay term, ensures that passengers would avoid congested airports which are otherwise attractive. The model is a useful tool for planning next generation air transportation systems as it enables the modeler to derive reasonable estimates of route and segment flows on a nationwide scale. The model also provides the analyst estimates of sensitivities of passengers to various factors, for example travel times and costs. The model could also be utilized to perform various "what-if" scenarios with the air transportation system. One of the shortcomings of any standard passenger choice dataset is that it has information only on the choice made by the passengers but has no information on the alternate choice set. In this model, a static alternate choice set (average route fares) was assumed. This assumption might not be realistic since fares fluctuate widely. Picking random fare values from the fare distribution might yield better results.

The model could possibly be enhanced by attempting a mixed logit for the O-D pairs with having only connecting itineraries by using a smaller calibration data set. However, care should be taken to ensure that the calibration dataset is not biased. The calibration could be performed for different calibration datasets to verify the stability of the mixed logit coefficients.

The coefficients for travel time and travel cost are both negative. This is expected as both increased travel time and fare result in increased disutility to the traveler. The coefficients with and without the inclusion of average airport delay in the travel time variable does not cause a significant change in the magnitude or the signs of the coefficients. This could be since, airport delays are still not very significant in most airports across the US, except for a few hubs such as ORD, EWR and LGA. However, as demand grows in the future and airport delays increase the flows obtained with and without the inclusion of airport delay could exhibit significant divergence. The coefficient for frequency is positive, as increased frequency leads to decreased schedule delay for the traveler and therefore, an increase in utility. The coefficients for seats is also positive since a greater number of seats offered produce a corresponding decrease in the probability of the passenger being denied a seat. The magnitude of the coefficient for seats is small as the number of seats offered per year on a leg is large. The coefficient for the number of legs is negative as passengers typically avoid connecting flights. However, the model leads to significant errors in the prediction of enplaning passengers for ATL and IAD (11% and 16% respectively). However, the introduction of dummy variables reduced the errors to 6% and 3% respectively. The model, which incorporates dummy variables, has a better log-likelihood ratio than the alternative model, which has no dummy variables. It can be inferred that the use of dummy variables increases the predictive power of the model. The t-statistics indicate that all variables are significant with a 95% level of confidence.

In Section 4.3 a family of models to derive the flight frequency and operating schedule in the Continental United States has been presented. Two approaches to the flight frequency problem have been presented. The first approach constructs a new “clean-slate” schedule disregarding the existing airline schedule. The second approach attempts to build a new schedule by modifying the existing schedule. The second approach is preferable as it produces better results and takes into account the current state of the airline industry. The model also accounts for physical constraints in the National Airspace System (NAS) such as aircraft payload range constraints and runway length constraints as these factors might preclude certain operating practices. The model also partly takes into account airline market dynamics, such as the agreements between the regional carriers and their mainline carriers. The model also provides a computationally efficient scheme to derive an integer solution from the output of the LP relaxation, by minimizing the degradation the optimal objective function. The model also takes into account stopover flights, which have not been addressed by most of existing literature on airline scheduling. Therefore, Formulation V which builds an incremental schedule with a variable fleet size is chosen for constructing the schedule for the future years.

The schedule generation module however does not address certain issues, which require more investigation and could be avenues for future research and greater enhancements to

the model. The main issue in the model is that the model does not take into account airline competition and does not distinguish between individual airlines. For example, in the segment from Cleveland, OH (CLE) and Detroit, MI (DTW) the model replaces the Boeing 737's and the Canadair CRJ-200 equipment by the Embraer 170, and similar aircraft as this is the most cost-effective solution. However, CLE-DTW segment is only served by Northwest and Continental Airlines. Both these airlines do not possess Embraer 170 or similar aircraft vehicles in their fleet as of 2004. The model solution could only happen if both Continental and Northwest cease operating the CLE-DTW market and another airline which possesses a large fleet of Embraer 170 or similar aircraft, for example JetBlue enters the CLE-DTW market which is not very probable. The model solution exhibits how the airline network would behave if all the airlines show cooperative behavior. In other words, each airline is willing to sacrifice market share and potential revenue in order to minimize cost for every player in the market. However, research by Hansen (1990) and Adler (2005) show that the airline business is fiercely competitive and airlines do not cooperate except in code-share markets. Unfortunately the game-theoretic approach proposed by the authors is only feasible for small or medium - sized networks. Nevertheless, discriminating individual airlines is an enhancement that could be an active area for future research. The model also does not take into account cost differences between airlines and assumes a common average value by aircraft type. For example, the operating cost for Southwest is on average about 2 cents per Available Seat Mile (ASM) lower than the mainline carriers.

The schedule generator module is a good first order approximation to study the state of the airline industry. Future enhancements could include modeling individual airlines and solving the schedule problem for each airline and obtaining airline competitive dynamics.

The network evolution module attempts to predict the future structure of the airline network by calibrating a statistical model based on the past trends and the current state of the NAS. The network evolution is a comprehensive modeling effort, as it predicts such real world phenomenon such as airline entry and exit from a market, development of new airports, development of new airline hubs and airline exit from current airline hubs. Most of the models discussed in Chapter 2, assume a static network structure. However, airline network topologies change constantly. The module also is computationally simple and does not employ complicated mathematical programming techniques making the model simpler to use in a strategic modeling exercise. The network module utilizes explanatory variables that are relatively simple to obtain, such as the operating cost per segment, load factor and the operating frequency. However, the network evolution module does not take into account airline competitive dynamics. An airline would be less inclined to cut a route on which the airline has a monopoly. On the other hand an airline would tend to withdraw from heavily competitive routes where the airline profit or market share is not significant and would tend to use the aircraft and crew resource freed to improve its market share and profit in another market which is less competitive. As in the case of the schedule generator module, the network evolver assumes that the airlines behave cooperatively and would collectively withdraw from a market that is not profitable. However, it is possible that a market would become profitable once the number of players in the market decreases and airlines would stop withdrawing from the market. In order to capture airline competitive dynamics, the network evolution module would have

to be enhanced by calibrating a separate probit model for each individual airline. The airline specific probit module could include airline competitive metrics, such as the Herfindahl Index or the airline market share. This airline specific model could then be applied to the NAS to determine, the particular airline that would cease operating in particular market and could potentially capture airline competition without requiring the use of a complex game theoretic formulation. The same enhancement could be applied to the other parts of the network evolution sub-module to increase the accuracy of the network evolution sub-module.

The output of the network evolution module predicts the probability of the route being removed and added. The output variable is the cumulative random normal values at the input variable. The coefficients for service frequency, profit per seat and load factor are negative which indicate that a higher value for these variables would make the input variable lower, and consequently lower the probability of the route being removed. The coefficient for the essential air service is also negative, which indicates that an airport that has been designated as an essential air service airport is less likely to lose service. On the other hand, the coefficient for distance is positive which indicates that long-haul routes are more likely to be cut by the airlines. At first glance, that seems counter-intuitive, however the distance variable is only significant for service from small and non-hubs, where the service is mainly connecting service to a hub rather than a point-to-point service. Increasing distance reduces the likelihood of a hub-to-spoke connection, as explained by Savage (2004), which validates the positive coefficient for distance. The regression equation that adds new hub-spoke links has positive coefficients for hub-to-spoke distance and the existence of a prior hub-to-spoke connection, which would indicate that the chances of a new hub-to-spoke connection decrease with increasing distance and the existence of a prior hub-spoke connection. The coefficient for the spoke demand and maximum spoke frequency were negative, which prove that both these variables have a positive effect on the likelihood of a new hub-spoke connection. Savage (2004) illustrates that spoke airports higher demand tends to have better hub connections.

The fare evolution sub-module predicts the macroscopic fare trends into the future. The macroscopic fare trends obtained are then applied to the individual route and market fares. The fare evolution module assumes that the macroscopic fare trends would be applicable on an individual market. However, individual market trends could be significantly different from the macroscopic nationwide trends due to airline competition, operating aircraft type and individual airline strategy. The fare trend sub-module, therefore only predicts the change of fares instead of the actual fares. The baseline values for the fares are the actual market fares from the DB1B. Actual fares could be obtained for each market by adopting a crude form of revenue management. The fare that gives the maximum revenue for each market could be used as the average market fare for the future years. This method requires a good estimation of demand elasticity with respect to fare for each market in the NAS. However, the demand model in TSAM could be utilized to obtain the demand elasticity and the optimal fare for each market. This methodology requires greater investigation and could be an active topic for future research.

The focus of the dissertation is to complete the supply side modeling in TSAM. TSAM has a detailed county-to-county demand model. However, TSAM did not have detailed supply side model to complement the demand model or the study the airline response to changes in demand. The dissertation also enables a modeler to capture the changes in fares, cost structure and fleet composition in the NAS. The dissertation cannot predict the market strategies of individual airlines. For examples, Southwest expanded in Pittsburgh International Airport (PIT) after 2004. This dissertation would treat Southwest, as a part of a larger “mega-airline” and cannot distinguish between service offered by United and Southwest. The fleet composition model used the dissertation is a predictor of airline behavior and cannot be used as a guide or a reference for airlines to buy aircraft. However if the dissertation is expanded to include individual airlines, the model results could be used to distinguish between service offered by individual airlines and could be used by airlines to make inventory decisions.

The model could be utilized by the FAA and other aviation planning organization to understand the evolution of the aviation system into the future. The model is driven by both supply and demand, and therefore provides a more reliable estimate than a model driven by pure demand or supply only. The FAA can employ the model to obtain reliable estimates of passenger enplanements, number of operations and consequently delays at major hubs. The FAA currently uses a model called the Terminal Area Forecast (TAF) to predict enplanements at the top 35 airports in the US. However, the TAF does not use a detailed supply side model in conjunction with demand. The TAF could be enhanced by applying the enplanements obtained from the model. Other government agencies, for examples the Joint Planning Development Office (JPDO), which are involved in planning for the NextGen air transportation system, could utilize the model to study the effects of new technologies in aviation on the aviation system. For example, if tiltrotor aircraft are introduced in the network in 2015, the model could predict the new markets that could be served by the new technology and the subsequent evolution of the system. Finally, the airlines could use the model to conduct market analysis and make inventory decisions.

Chapter 7 References

- Adler, N., Hub-Spoke Network Choice under Competition with an Application to Western Europe. *Transportation science*, 2005. **39**(1): p. 58-72.
- Adler, N., Competition in a Deregulated Airline Market. *European Journal of operations research*, 2001. **129**(2): p. 337-345.
- Aldregghi, M., et al., Network Competition - The Coexistence of Hub-And-Spoke And Point-To-Point Systems. *Journal of Air Transport Management*, 2005. **11**(5): p. 328-334.
- Allenby, Greg. M and Rossi, Peter. E, There Is No Aggregate Bias: Why Macro Logit Models Work, *Journal of Business and Economic Statistics*, 1991, **9**(1), p. 1-14.
- Aykin, T., Networking Policies for Hub-And-Spoke Systems with Application to the Air Transportation System. *Transportation science*, 1995. **29**(3): p. 201-221.
- Barnhart, C. and M. Lohatepanont, Itinerary Based Airline Fleet Assignment. *Transportation science*, 2002. **36**(2): p. 199-217.
- Bauer, P., Airline Hubs: A Study of Determining Factors and Effects. *Economic review - Federal reserve bank of Cleveland*, 1987. **1987 Quarter 4**: p. 13-19.
- Bazaraa, M.S, J.J Jarvis and H.D Sherali, *Linear Programming and Network Flows*. 3rd Edition, John Wiley and Sons Inc., New York, NY, 2005.
- Braanlund, Railway Timetabling Using Lagrangian Relaxation. *Transportation science*, 1998. **32**(4): p. 358-367.
- Coldren, G.M., F.S. Koppelman, K. Kasturirangan, A. Mukherjee, Air Travel Itinerary Share Prediction: Logit Model Development At A Major U.S. Airline, *Journal of Air Transportation Management*, 2003, **9**(6), p. 361-369.
- Coldren, G.M. and F.S. Koppelman, Modeling the Competition Among Air-Travel Itinerary Shares: GEV Model Development, *Transportation Research Part A*, 2005, **39**(5),p. 345-365
- Dennis, N., Scheduling Issues and Network Strategies for International Airline Alliance,. *Journal of Air Transport Management*, 2000. **6**: p. 75-85.
- Desaulniers, G., J. Desrosiers, and Y. Dumas, *Daily Aircraft Routing and Scheduling. Management Science*, 1997. **43**(6): p. 841-855.
- Dobson, G. and P.J. Lederer, Airline Scheduling and Routing in a Hub-and-Spoke System. *Transportation science*, 1993. **27**(3): p. 281-297.
- Erdmann, A., A. Nolte, and R. Noltemeier, Modeling and Solving the Airline Schedule Generation Problem. 1999, University of Cologne. p. 1-23.
- Feo, T. and J. Bard, Flight Scheduling and Maintenance Base Planning. *Management Science*, 1989. **35**(12): p. 1415-1432.
- Ghobrial, A. and A. Kanafani, Future of Airline Hubbed Networks: Some Policy Implications. *Journal of Transportation Engineering*, 1995. **121**(2): p. 124-133.
- Gillen, D. and W. Morrison, Regulation, Competition and Network Evolution in Aviation *Journal of Air Transport Management*, 2005. **11**(3): p. 161-174.
- Glover F, Management Decision and Integer Programming, *The Accounting Review*, 1969, **44**(2): p.300-303
- Gosling, G. and M. Hansen, Projects for Increasing Aircraft Size at Congested Airports. *Transportation Research Record*, 1995. **1744**: p. 1-8.
- Hansen, M., Airline Competition in a Hub Dominated Environment: An Application of

- Non-Cooperative Game Theory. *Transportation Research part B*, 1990. **24**(1): p. 27-43.
- Heidemann, D. Note on the Entropy Maximization Principle in Traffic and *Transportation Studies*, 2000. Beijing, People's Republic of China.
- Hendricks, K., M. Piccione, and G. Tan, Equilibria in Networks. *Econometrica*, 1999. **67**(6): p. 1407-1434.
- Hong, S. and P.T. Parker, Air Traffic Network Equilibrium: Towards Frequency, Price and Slot Priority Analysis. *Transportation Research part B*, 1992. **26**(4): p. 307-323.
- Hsiao, C.-Y. and M. Hansen. Air Transport Network Flows: An Equilibrium Model. In *Proceedings of the 84th Annual Meetings of the Transportation Research Board*. 2004. Washington, DC: National Academy of Sciences.
- Hsu, C.-I. and Y.-H. Wen, Determining Flight Frequencies on an Airline Network With Supply-Demand Interactions. *Transportation Research part E*, 2003. **39**(6): p. 417-441.
- Jalliet, P.S.S., Airline Network Design and Hub Location Problems. *Location Science*, 1996. **4**(3): p. 195-212.
- Jansson, K. and B. Ridderstolpe, A Method for the Route Choice Problem in Public Transport Problem in Public Transport Systems. *Transportation science*, 1992. **26**(3): p. 246-251.
- Kanafani, A. and A. Ghobrial, Airline Hubbing- Some Implications for Airport Economics. *Transportation Research part A*, 1984. **19**(1): p. 15-27.
- Lederer, P. and R. Nambimadom, Airline Network Design. *Operations research*, 1998. **46**(6): p. 785-804.
- Lee, D., Concentration and Price Trends in the US Airline Industry: 1990-2000. *Journal of Air Transport Management*, 2003. **9**(2): p. 91-101.
- Lee, D. and H. Ito, Domestic Code Sharing Practices in the US Airline Industry. *Journal of Air Transport Management*, 2005. **11**: p. 89-97.
- Lohatepanont, M. and C. Barnhart, Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment. *Transportation science*, 2004. **38**(1): p. 19-32.
- Marianov, V., D. Serra, and C. ReVelle, Location of Hubs in a Competitive Environment. *European Journal of operations research*, 1999. **114**(2): p. 363-371.
- Morton, O.K., A Quadratic Integer Program for the Location of Interacting Hub Facilities. *European Journal of operations research*, 1987. **32**: p. 393-404.
- Morton, O.K., The Location of Interacting Hub Facilities. *Transportation science*, 1986. **20**(2): p. 92-105.
- Niedringhaus, W. P. (2004), The Jet Wise Model of National Airspace System Evolution, *Simulation*, **80**(1), p. 45-58.
- Oum, T.H., A. Zhang, and Y. Zhang, Airline Network Rivalry. *Canadian Journal of Economics*, 1995. **28**(4a): p. 836-857.
- Phillips, R. and D. Boyd, An Algorithm for Calculating Consistent Itinerary Flows. *Transportation science*, 1991. **25**(3): p. 225-239.
- Rappaport A, Sensitivity Analysis in Decision Making, *The Accounting Review*, 1967 **42**(3): p.441-456.
- Rietveld, P. and M. Brons, Quality of Hub and Spoke Networks: The Effects of

- Timetable Coordination on Waiting Time and Rescheduling Time. *Journal of Air Transport Management*, 2001. **7**: p. 241-249.
- Rosenberger, J., E. Johnson, and G. Nemhauser, A Robust Fleet-Assignment Model with Hub Isolation and Short Cycles. *Transportation science*, 2004. **38**(4): p. 357-368.
- Savage, I. and B. Scott, Deploying Regional Jets to Add New Spokes to a Hub. *Journal of Air Transport Management*, 2004. **10**(2): p. 147-150.
- Sherali, H.D., Bae, K. and Haouari, M. Polyhyderal Analysis and a Benders Decomposition Approach for an Integrated Schedule Planning and Fleet Assignment Problem. Manuscript, Dept of Industrial and System Engineering, Virginia Tech, Blacksburg, VA, 2009.
- Sherali, H.D., Manuscript, Dept of Industrial and System Engineering, Virginia Tech, Blacksburg, VA, 2008.
- Simpson, R.W. Constructive Approach to Route System Planning. In *Proceedings of AGIFORS conference*. 1969. Broadway, England.
- Soumis, F. and A. Nagurney, A Stochastic Multiclass Network Equilibrium Model. *Operations research*, 1992. **41**(4): p. 721-730.
- Stojkovic, G., F. Soumis, and J. Desrosiers, An Optimization Model for a Real-Time Flight Scheduling Problem. *Transportation Research part A*, 2002. **36**(9): p. 779-788.
- Stojkovic, M. and F. Soumis, An Optimization Model for the Simultaneous Operational Flight and Pilot Scheduling Problem. *Management Science*, 2001. **47**(9): p. 1290-1305.
- Swan, W., Airline Route Developments: A Review of History. *Journal of Air Transport Management*, 2002. **8**: p. 349-353.
- Teodorovic, D., Multicriteria Model to Determine Flight Frequencies on an Airline Network Under Competitive Conditions. *Transportation science*, 1989. **23**(1): p. 14-24.
- Teodorovic, D., Multiattribute Aircraft Choice for an Airline Network. *Journal of Transportation Engineering*, 1987. **112**(6): p. 634-645.
- Teodorovic, D., Flight Frequency Determination. *Journal of Transportation Engineering*, 1983. **109**(5): p. 747-757.
- Teodorovic, D. and M. Kalic, A Fuzzy Route Choice Model for Air Transportation Networks. *Transportation planning and technology*, 1994. **19**: p. 109-119.
- Teodorovic, D. and G. Stojkovic, Model for Operational Daily Airline Daily Scheduling. *Transportation planning and technology*, 1990. **14**: p. 273-285.
- Van Zuylen, H. and L. Willumsen, The Most Likely Trip Matrix Estimated from Traffic Counts. *Transportation Research part B*, 1979. **14**: p. 281-293.
- VanAerde, M. and H. Rakha. Comparison of Static Maximum Likelihood Origin-Destination Formulations. in *Transportation and Traffic Theory: Flow, Dynamics and Human Interaction, Proceedings of the 16th International Symposium on Transportation and Traffic Theory (ISTTT16)*,. 2005.
- VanAerde, M. and H. Rakha. Estimations of O-D Matrices: The relationship between practical and theoretical considerations. *Proceedings of the 82nd Annual meeting of the Transportation Research Board*. 2003. Washington,DC: National Academy of Sciences.
- Veldhuis, J., The Competitive Position of Airline Networks. *Journal of Air Transport*

- Management*, 1997. **3**(4): p. 181-188
- Vovsha, P. and S. Bekhor, Link-Nested Logit Model of Route Choice: Overcoming the Route Overlapping Problem. *Transportation Research Record*, 1998. **1645**: p. 133-141.
- Weber, M. and G. Williams, Drivers of Long-Haul Air Transport Development. *Journal of transport geography*, 2001. **9**(4): p. 243-254.
- Wei, W. and M. Hansen, Impact of Aircraft Size and Seat Availability on Airlines Demand and Market Share in Duopoly Markets. *Transportation Research part E*, 2005. **41**(4): p. 315-327.
- Wiedner, T., Hubbing in the US Air Transportation System: An Econometric Approach. *Transportation Research Record*, 1995. **1562**: p. 28-36.
- Yan, S., The Planning of Aircraft Routes and Flight Frequencies in Airline Network Operations. *Journal of advanced transportation*, 1996. **35**(1): p. 33-46.
- Yan, S., C.-H. Lai, and H.-C. Chia, A Short Term Flight Scheduling Model for International Express Package Delivery. *Journal of Air Transport Management*, 2005. **11**(6): p. 368-374.
- Yan, S. and P. Tu Yu, Multifleet Routing and Multistop Flight Scheduling for Schedule Perturbation. *European Journal of operations research*, 1997. **10**
- Bureau of Transportation Statistics,
http://www.bts.gov/press_releases/2005/bts052_05/html/bts052_05.html, 2003
- Wikipedia, http://en.wikipedia.org/wiki/Essential_Air_Service, 2008
- Wikipedia, <http://en.wikipedia.org/wiki/PIT> , Pittsburgh International Airport, 2008
- Wikipedia, <http://en.wikipedia.org/wiki/MDW> , Chicago Midway Airport, 2008.