A DESIGN METHODOLOGY FOR EVOLUTIONARY AIR TRANSPORTATION NETWORKS

A Thesis
Presented to
The Academic Faculty

by

Eunsuk Yang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Aerospace Engineering

Georgia Institute of Technology
August 2009

A DESIGN METHODOLOGY FOR EVOLUTIONARY AIR TRANSPORTATION NETWORKS

Approved by:

Dr. Dimitri N. Mavris,
Committee Chair
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Hojong Baik
Civil, Architectural and
Environmental Engineering
Missouri University of Science and Technology

Dr. Jung-Ho Lewe
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Daniel DeLaurentis
School of Aeronautics and
Astronautics
Purdue University

Dr. Daniel Schrage
School of Aerospace Engineering
Georgia Institute of Technology

Date Approved: May 15th, 2009

To my family and friends,

With all my love and respect.

ACKNOWLEDGEMENTS

I would like to give my first thank you to Dr. Mavris, my advisor, who finds joy in giving opportunities to people and nurturing them. I also became one of the students who believes that he is more a mentor than just an advisor. I would like to extend my gratitude to Dr. Jung-Ho Lewe for his guidance throughout this research. To Dr. Daniel DeLaurentis for his kind advice and for the mentorship during my initial years at Georgia Tech. To Dr. Daniel Schrage and Dr. Hojong Baik for their continuous support and sharing their wisdom and knowledge.

Special thanks goes to Samson who helped me in many ways while working in the same team. Also, I would like to show my appreciation to Chirag, Eric, Henry, Ramachandra, Ruth, Samer, Taewoo, and Yuan for their friendship and support.

I would also like to thank Dr. Latham who was always able to find time for me whenever I had problems with database server and other computer related problems.

I would not have come this far without you.

Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS OR ABBREVIATIONS

ACES Airspace Concepts Evaluation System.

ATM Air Traffic Management.

ATS American Travel Survey.

FAA The Federal Aviation Administration.

JPDO Joint Planning & Development Office.

Mi Mi Model.

NAS National Airspace System.

NASA National Aeronautics and Space Administration.

NTS National Transportation System.

SATS Small Aircraft Transportation System.

TAF Terminal Area Forecast.

τ_{ij} Origin-Destination demand.

t_{ij} True Origin-Destination Demand.

TSAM Transportation System Analysis Model.

SUMMARY

The air transportation demand at large hubs in the U.S. is anticipated to double in the near future. Current runway construction plans at selected airports can relieve some capacity and delay problems, but many are doubtful that this solution is sufficient to accommodate the anticipated demand growth in the National Airspace System (NAS). With the worsening congestion problem, it is imperative to seek alternative solutions other than costly runway constructions. In this respect, many researchers and organizations have been building models and performing analyses of the NAS. However, the complexity and size of the problem results in an overwhelming task for transportation system modelers. This research seeks to compose an active design algorithm for an evolutionary airline network model so as to include network specific control properties. An airline network designer, referred to as a network architect, can use this tool to assess the possibilities of gaining more capacity by changing the network configuration.

Since the Airline Deregulation Act of 1978, the airline service network has evolved into a distinct Hub-and-Spoke (H&S) network. Enplanement demand on the H&S network is the sum of Origin-Destination (O-D) demand and transfer demand. Even though the flight or enplanement demand is a function of O-D demand and passenger routings on the airline network, the distinction between enplanement and O-D demand is not often made. Instead, many demand forecast practices in current days are based on scale-ups from the enplanements, which include the demand to and from transferring network hubs. Based on this research, it was found that the current demand prediction practice can be improved by dissecting enplanements further into smaller pieces of information. As a result, enplanement demand is decomposed into

intrinsic and variable parts. The proposed intrinsic demand model is based on the concept of 'true' O-D demand which includes the direction of each round trip travel. The result from using true O-D concept reveals the socioeconomic functional roles of airports on the network. Linear trends are observed for both the produced and attracted demand from the data. Therefore, this approach is expected to provide more accurate prediction capability.

With the intrinsic demand model in place, the variable part of the demand is modeled on an air transportation network model, which is built with accelerated evolution scheme. The accelerated evolution scheme was introduced to view the air transportation network as an evolutionary one instead of a parametric one. The network model takes in intrinsic demand data before undergoing an evolution path to generate a target network. The results from the network model suggest that air transportation networks can be modeled using evolutionary structure and it was possible to generate a closely emulated NAS. A dehubbing scenario study of Lambert-St. Louis International Airport demonstrated the prediction capability of the proposed network model. The overall process from intrinsic demand modeling and evolutionary network modeling is unique and it is highly beneficial for simulating active control of air transportation networks.

CHAPTER I

INTRODUCTION

1.1 Motivation

1.1.1 Growing air transportation demand

Air transportation systems have become a crucial part in our daily life. The society's mobility freedom has grown dramatically since air transportation became an affordable mode of transportation and its supporting infrastructure is well established. In particular, air travel made it possible for the majority of people to go to places that was too expensive or too tedious to reach previously. In modern days, air travel has become the dominant travel mode in long-haul travel, whereas automobiles remain the dominant travel mode in short-haul travel because of its convenience, low cost, and high accessibility. As demand for air travel grows, congestion issues are becoming more and more severe because the service capacity is not matching up to the growth. The Federal Aviation Administration (2008*a*) (FAA) has projected that the revenue passenger enplanements[1] will grow from 776.5 million in 2008 to 1,292.9 million in 2025 as shown in Table 1. It is also predicted that both the passenger and freight demand will double in the next 20 years and triple within 50 years (The Federal Transportation Advisory Group, 2001).

With an increasing demand and thus worsening congestion in the National Airspace System (NAS), it is imperative to seek alternative solutions other than costly runway constructions. Some adjustments to the NAS can relieve capacity and delay problems, which include flight schedule shifts to non-congested periods of the day, improvement

[1]Enplanements (or Enplaned passengers): number of passengers who have boarded any commercial flights

Table 1: Total scheduled U.S. passenger traffic. [Source: Federal Aviation Administration (2008*a*)]

Fiscal Year	Revenue Enplanements (Millions)			Revenue Pax Miles (Billions)		
	DOMESTIC	I-NATIONAL	SYSTEM	DOMESTIC	I-NATIONAL	SYSTEM
Historical*						
2000	641.2	56.4	697.6	512.8	181.8	694.6
2001	625.8	56.7	682.5	507.9	183.3	691.1
2002	575.1	51.2	626.3	473.4	158.2	631.6
2003	587.8	53.3	641.2	492.7	155.6	648.3
2004	628.5	60.5	689.0	540.2	177.0	717.2
2005	669.4	67.4	736.9	576.9	197.2	774.2
2006	668.4	71.6	740.0	582.4	208.5	790.9
2007E	689.4	75.3	764.7	600.1	221.2	821.4
Forecast						
2008	696.2	80.2	776.5	603.8	241.0	844.8
2009	720.6	85.0	805.7	624.5	259.9	884.4
2010	746.2	89.8	836.0	650.2	278.0	928.3
2011	767.0	94.7	861.7	671.7	296.2	967.9
2012	789.4	99.3	888.7	696.6	312.6	1,009.2
2013	812.6	104.0	916.6	722.8	329.6	1,052.4
2014	834.6	109.0	943.6	747.3	347.4	1,094.7
2015	859.0	114.1	973.1	776.2	365.9	1,142.1
2016	882.4	119.4	1,001.8	803.8	385.2	1,189.0
2017	907.8	125.0	1,032.8	835.6	405.5	1,241.1
2018	933.5	130.9	1,064.4	867.4	426.5	1,293.9
2019	958.6	136.6	1,095.2	900.8	447.2	1,347.9
2020	984.0	142.5	1,126.5	935.2	468.7	1,403.8
2021	1,008.9	148.2	1,157.2	969.6	489.6	1,459.2
2022	1,035.0	154.4	1,189.4	1,006.3	512.3	1,518.6
2023	1,061.8	160.8	1,222.6	1,044.6	535.9	1,580.5
2024	1,089.3	167.5	1,256.8	1,084.7	560.4	1,645.1
2025	1,118.2	174.7	1,292.9	1,126.9	586.8	1,713.7
Avg Annual						
Growth: 2000-07	1.0%	4.2%	1.3%	2.3%	2.8%	2.4%
2007-10	2.7%	6.1%	3.0%	2.7%	7.9%	4.2%
2010-20	2.8%	4.7%	3.0%	3.7%	5.4%	4.2%
2007-25	2.7%	4.8%	3.0%	3.6%	5.6%	4.2%

of air traffic control, and NAS modernization. But it is doubtful whether those adjustments are sufficient for a doubled or tripled demand (Hunter et al., 2005). In this respect, many researchers and organizations have been building models and performing analysis of the NAS (O'Kelly and Miller, 1994; Bania et al., 1998; Bonnefoy and Hansman, 2007).

As far as dealing with the capacity and delay problems of the NAS, the most commonly used metrics are the numbers of operations and enplanements at an airport level. A typical mental model behind this approach is depicted in Figure 1, demonstrating the relationships between directly observable top-level metrics for a particular airport.

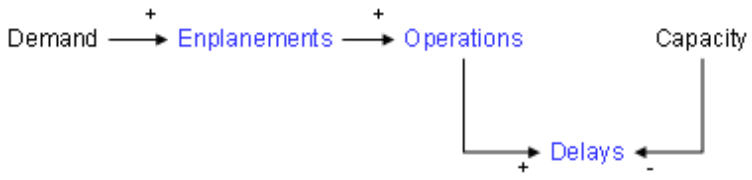


Figure 1: Enplanements, operations, and delay at an airport. [Source: Lewe (2008)]

In Figure 1, travel demand is directly translated into enplanements. For higher enplanements, operations are higher when other conditions remain the same. Also, higher operation increases delays ((+) sign) while higher capacity decreases delays ((-) sign), assuming everything else stays equal. Also, note that this would apply to point-to-point (P2P) networks as well, where delays are additive and every itinerary has one segmental trip. When an air transportation network is not purely P2P, this reasoning does not hold.

1.1.2 Evolving air transportation supply

One of features that make the air transportation system distinguished from other transportation is that the system has experienced topological reconfiguration in its

nature of network. Since the Airline Deregulation Act (ADA) of 1978 the NAS service network has evolved into a so-called hub-and-spoke (H&S) system (U.S. General Accounting Office, 1996; Poole and Butler, 1999). Figure 2 shows the route maps of Delta airlines, illustrating the evolution path.

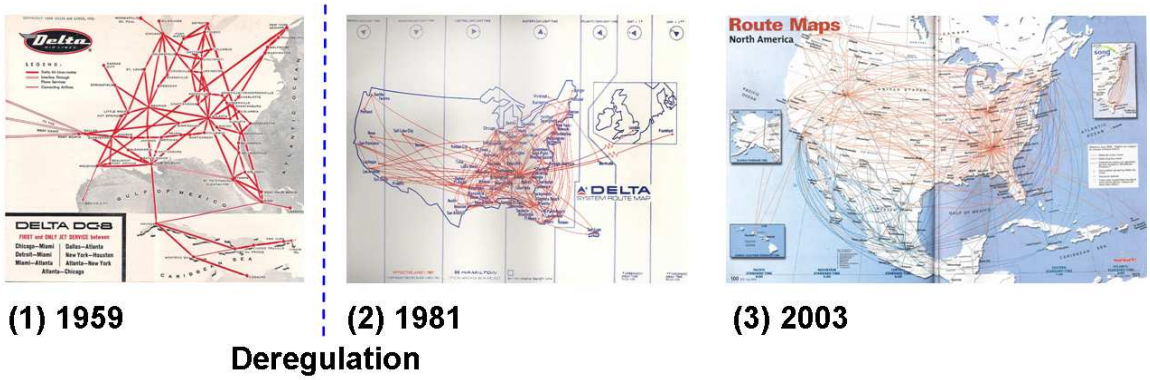


Figure 2: Historic Delta Airline route maps

A discontinuity observed between 1959 and 1981 is largely credited to the ADA. In 1975, entry restriction was eased by the Civil Aeronautics Board (CAB). The ADA was issued in 1978 and by 1983 no control was enforced. Air carriers, in an attempt either to maximize their profit or just to simply survive, have adopted the H&S system since then.[2] Relatively low demands at spoke airports are aggregated at hubs and bigger airplanes are used at higher frequency between big hub airports. Major carriers generally service larger cities where demand is high and utilize large commercial aircraft. Major carriers have been operating with regional carriers to open the small community markets that they could not service before with their large aircraft. The technical and financial help from the major carriers helped the growth of regional carriers, widening the air transportation network coverage. In many cases, regional airlines partnered with major carriers – called code-sharing – take care of the smaller demand of hub-to-spoke connections. This way, major airlines can expand

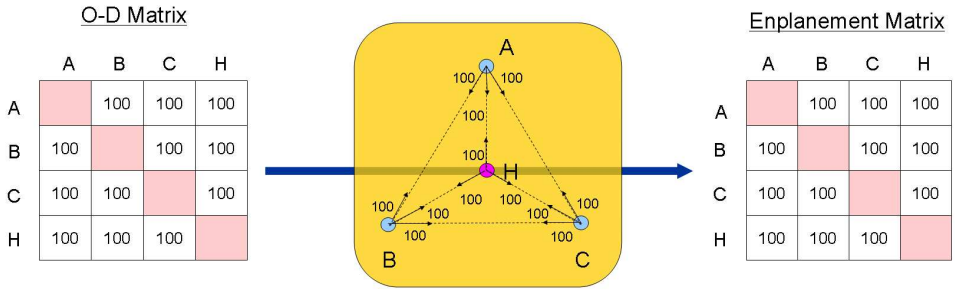
[2]Further information about economic aspects of airline H&S structure and airline cost with economies of density and scale can be found in Button (2002) and Caves et al. (1984).

their service network and utilize economies of scale at the same time. Travelers gain benefits from the H&S system as well. The travel demand from a small city to a big hub city is generally larger than the demand from a small city to other small ones. Also, there is good established demand between hub cities. Therefore, travelers also have the benefit of having more frequent departures with H&S system. Because of these benefits, widespread adoption of H&S after deregulation was also observed in Europe (Burghouwt and Hakfoort, 2001; Burghouwt et al., 2003).

1.1.3 Network effects: concentration and propagation

Even though the H&S system is a valuable choice for air carriers, it increases traffic on the NAS. For hub airports, enplanements are not only from their own Origin-Destination (O-D) demand but also from transferring passengers. To show the effects of hub concentration on enplanements and operations, an example below was devised in Figure 3. It shows three different cases of the network usage to fulfill the same O-D travel demand on a transportation network. O-D demand is shown on the left matrix and enplanements on the example network is shown on the right matrix. The network diagram in the center shows enplanements on each segment.

Figure 3a shows the P2P network, where the O-D demand is the same as enplanements demand on the network. O-D demand, which is different from true O-D demand to be introduced later, is a symmetric demand where the incoming demand equals the outgoing demand. Figure 3b shows the case when 50 percent of the demand between A, B, and C have to go through the hub airport H. As a result, the enplanements at H should increase while the enplanements at other spoke airports remain the same. In this specific case, the total enplanements on the network increase by 25 percent and the operations decrease by 25 percent if 10 passenger aircrafts are used between spoke airports and 50 passenger aircrafts are used for hub connections. Figure 3c shows the case where all the traffic between A, B, and C have to be



	A	B	C	H	Sum
Enplanements	300	300	300	300	**1200**
Operations	22	22	22	6	**72**

Assumption: 10 pax A/C btw spokes, 50 pax A/C for hub connection

(a) Perfect P2P case

	A	B	C	H	Sum	%Diff
Enplanements	300	300	300	600	**1500**	**+25%**
Operations	14	14	14	12	**54**	**-25%**

Assumption: 10 pax A/C btw spokes, 50 pax A/C for hub connection

(b) H&S emerges

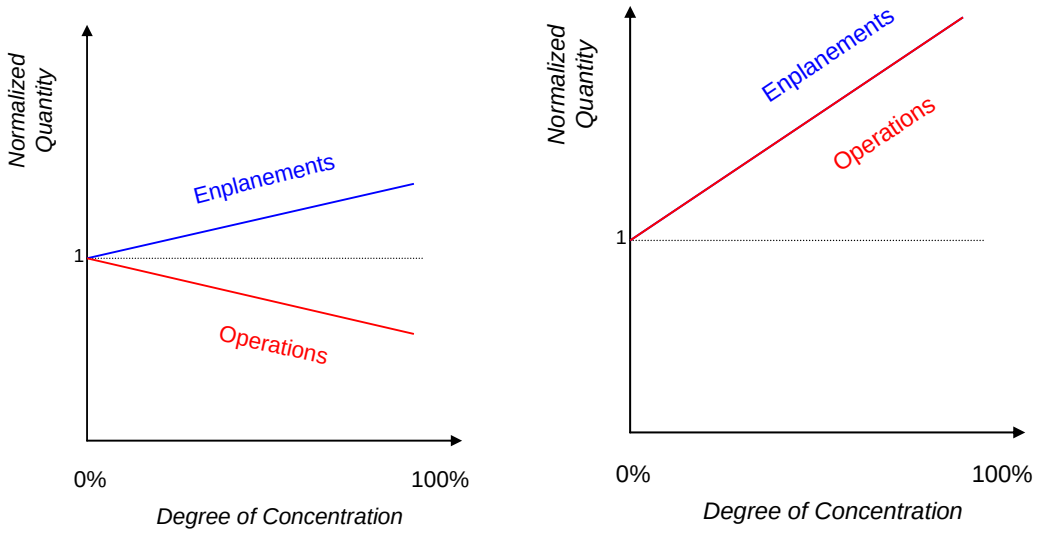
	A	B	C	H	Sum	%Diff
Enplanements	300	300	300	900	**1800**	**+50%**
Operations	6	6	6	18	**36**	**-50%**

Assumption: 10 pax A/C btw spokes, 50 pax A/C for hub connection

(c) Perfect H&S case

Figure 3: Enplanements and operations change for selected network usage

made through the hub airport H. In this extreme case, the total enplanements on the network increase by a dramatic 50 percent (from 1200 to 1800) and the operations decrease by 50 percent (from 72 to 36). As can be seen in this example, enplanements or operations depend heavily on the extent of hubbing degree. In all of these cases, the true origin-destination demands are the same. The only difference is how the network transfers passengers. As a summary, the H&S system increases enplanements and reduces operations at the system level(Figure 4a). However, both enplanements and operations increase at the hub airport compared to the P2P case (Figure 4b). Horizontal axes of the figure represents hub concentration of zero percent (P2P case) to 100 percent (full H&S case) and vertical axes is notional normalized quantity with 1 to designate the situation of P2P case.



(a) System (b) Hubs

Figure 4: H&S effect on enplanements and operations at system and hubs. [Source: Lewe (2008)]

The current NAS lies in between the perfect P2P and the perfect H&S system. It is a highly concentrated service network, as can be seen in Table 2. Note that about 97 percent of enplanements occur at 137 hub airports. Enplanements by each airport type is shown in Table 3. These tables show that demand at hub airports can be a

bottle-neck for the overall performance of the NAS.

Table 2: U.S. airport categorization by Federal Aviation Administration in 2005

Airport Type	Definition	Number	Percentage of Enplanements
Large Hub	1 % or more	30	69.04
Medium Hub	At least 0.25 %, but less than 1 %	38	20.02
Small Hub	At least 0.05%, but less than 0.25%	69	7.75
All Hub		137	96.81
Nonhub Primary[3]	More than 10,000, but less than 0.05%	247	3.10
All Primary		384	99.91
Non-Primary CS[4]	At least 2,500 and no more than than 10,000	130	0.09
All CS		514	100.00

Table 3: Enplanements by airport hub types, 2001-2005. [Data source: FAA Passenger and All-Cargo Statistics]

Hub Type	2001	2002	2003	2004	2005
Large	457,147,860	447,105,714	464,486,847	484,948,605	508,197,766
Medium	130,637,640	127,082,379	115,177,169	141,078,743	147,411,269
Small	50,350,109	48,664,326	50,202,980	57,569,857	56,246,103
Nonhub	21,287,219	20,286,334	20,178,352	21,196,521	22,826,796
Total	659,422,828	643,138,753	650,045,348	704,793,726	734,681,934

In summary, travel demands are met by the airlines that utilize evolving H&S system which generates ever-changing "network effects". It is important to note that the portion of the network-induced delays become significant when hub concentration increases. In this light, Figure 1 misses an important component of the dynamics between enplanements, operations, and delays. Considering the network effects, the realized relationships on the air transportation network can be depicted as in Figure 5.

Delay problems become worse when hub concentration increases since the high flight demand at hubs becomes bottlenecked for the connected trips. Therefore, delays at hubs are not isolated but instead propagates throughout the network. Therefore, the basis of the present research should be two-fold: First, intrinsic demand modeling

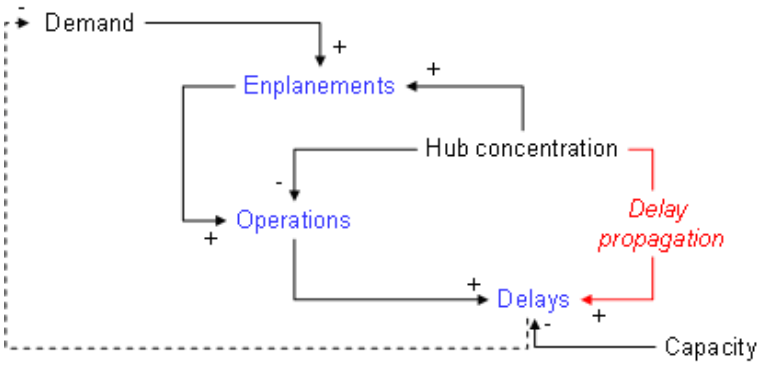


Figure 5: Enplanements, operations, and delay on H&S system. [Source: Lewe (2008)]

and second, network effect modeling. Intrinsic demand should be predicted prior to determining enplanements and operations in the network model, discussed next.

1.2 Research objectives

1.2.1 Characterization of intrinsic demand

Even though enplanements are a function of network usage, many of the current practices of the NAS demand estimation are scale-ups from the baseline as exemplified by Long, Lee, Gaier, Johnson and Kostiuk (1999). Even the Next Generation Air Transportation System (NGATS) – the most recent NAS study being conducted by the Joint Planning & Development Office (JPDO) – has its demand prediction based on scale-ups from TAF (Borener et al., 2006). Figure 6 shows baseline future scenarios for JPDO evaluation and analysis. As can be seen in the graph, baseline TAF projection is scaled-up with a higher growth rate for the future scenario generation.

In many works such as Schleicher et al. (2007), 2X- and 3X-demand scenarios have often been discussed but not defined clearly. Further, Bhadra et al. (2003) and Bhadra (2003) also pointed out that operations on the NAS is made up of specific airport-to-airport flows, which can have different behavior compared to growth of terminal area demand. Scale-up prediction does not generate much problems when the NAS is dominated by a P2P network. But as the NAS converts into a more H&S system,

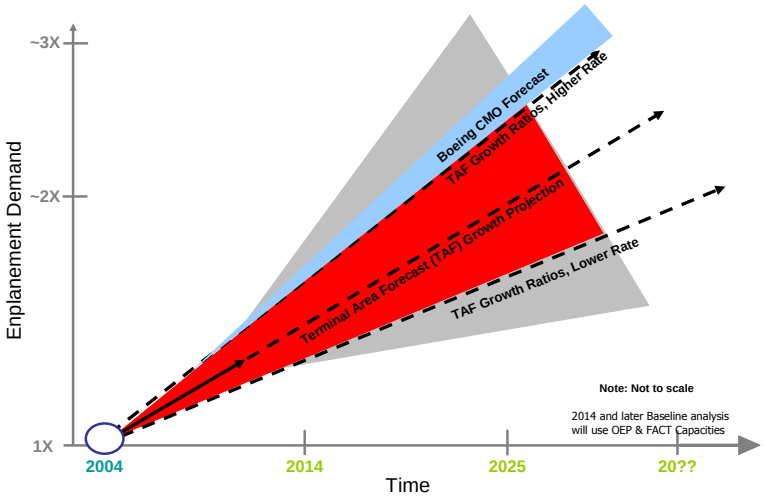


Figure 6: JPDO future demand projections. [Source: Borener et al. (2006)]

issues that were not important for P2P become prominent. The network load depends on the routings of the passengers. Therefore, the study to reveal the intrinsic demand of air transportation should proceed any other studies and the transfer demand should be modeled separately.

In this research, the demand is decomposed into intrinsic and variable ones and a concept of true O-D demand is introduced to include the directional aspect of round trip demands. Only round trips are considered for this research because most of the air travel is round trip due to the difficulty of using other modes of transportation for long distance travel. The relationships between the true O-D, O-D, and enplanements demands are represented in Figure 7. True O-D demand is obtained by recording only the outbound portion of each itinerary (leftmost matrix in Figure 7). This matrix is important since it represents pure trip generation (row) and attraction (column) factors of each airport. If the diagonal elements are summed together – i.e, M(a,b) + M(b,a) – the matrix becomes an O-D demand matrix. An O-D demand is the demand that needs to be transported. Enplanements numbers realized on the NAS depends on the topological usages of the airline network. For example, O-D demands (center matrix in Figure 7) can be serviced on a H&S network to generate enplanements on

the network (rightmost matrix in Figure 7). This example corresponds to the case reviewed in Figure 3b on page 6. Therefore, true O-D demand is one of the defining measures for the NAS along with enplanements characteristics.

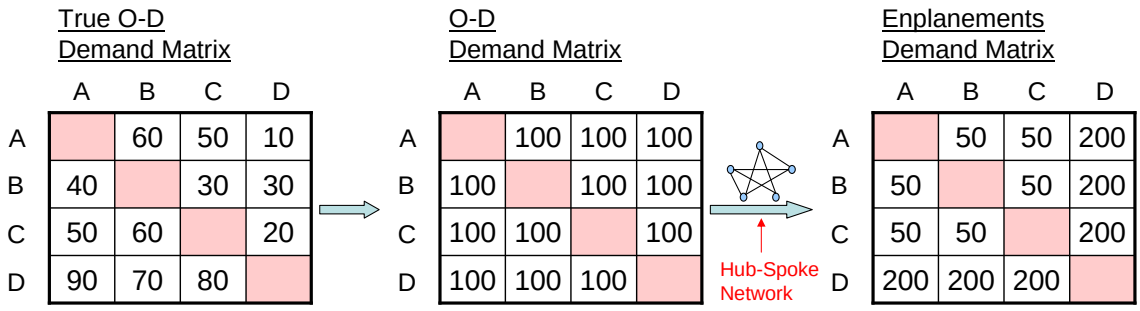


Figure 7: Three types of demand: True O-D, O-D, enplanements demand

1.2.2 Active formulation of network

To recall, the post-deregulation evolution of the NAS happened in a passive manner, meaning out of the control of the aviation authorities, only governed by natural market forces created between consumers (demand-side) and service providers (supply-side). As demand continues to outgrow the pace of infrastructure expansion, the lack of control and understanding of the NAS has resulted in a heavily depended air transportation network that may not be capable of handling the expected demand growth in the near future. Despite so, there has been little effort in the active design of the air transport network with predictive capabilities. These capabilities are imperative as are design methodologies that can be incorporated to guide the much needed reshaping of the NAS. To stress the need for such an active design tool, the notion of the transportation architect is introduced, i.e. the engineers and decision makers, who are responsible for formulating, analyzing, and optimizing the future network concepts. To perform his/her tasks, the architect requires tools that offers the capability to predict performance of the air transportation network as control and condition variables change. It is also important to be able to reflect these changes

effectively to the design process of the transportation network. The architect would need a parametric tool that can evaluate various scenarios with minimal setup time. Therefore, the modeling philosophy of this research is to build a comparatively simple model that captures system-level dynamics. It is postulated that evolution is a key underlying behavior in airline networks that enables practical simulation of the network.

The process of getting knowledge on a system is evolutionary in nature. Table 4 lists eight stages of knowledge that is defined by Bohn (1994). According to the definition, the status of the research on the NAS demand and network properties is at the level between *3* and *5*. This reasoning was made because the effect of modeling evolution in airline network modeling has not been tried and thus not clearly understood. The main goal of this research is to set up a framework for a demand- and network-centric analysis of the air transportation system and to expand the state-of-the-art airline network modeling technique to include evolutionary behaviors.

Table 4: Stages of knowledge. [Source: Bohn (1994)]

Stage	Name	Comment	Typical Form of Knowledge
1	Complete ignorance		Nowhere
2	Awareness	Pure art	Tacit
3	Measure	Pretechnological	Written
4	Control of the mean	Scientific method feasible	Written and embodied in hardware
5	Process capability	Local recipe	Hardware and operating manual
6	Process characterization	Tradeoffs to reduce costs	Empirical Equations (numerical)
7	Know why	Science	Scientific formulas and algorithms
8	Complete knowledge	Nirvana	

1.3 Organization of the thesis

This thesis starts with an introduction to the overall research field with identification of the problems in Chapter 1. In Chapter 2, existing models and theories that are related to demand prediction and network simulation are reviewed. Research questions and hypothesis to achieve the objective of this research will be identified in this chapter following the literature search. Chapter 3 describes general ideas and approaches taken to solve the problem and the proposed environment is presented. In Chapter 4, ideas and approaches are turned into concrete models. Detailed parameter settings for the models are also determined. In Chapter 5, a theoretical baseline that captures general behavior of the NAS is determined. Also, a practical baseline is determined that emulates the NAS closely. With the practical baseline, a scenario study of dehubbing at Lambert-St. Louis International Airport (STL) is conducted. Finally, conclusions and recommendations for future research are summarized in Chapter 6. Also, the contributions of this research to the existing knowledge base is addressed.

CHAPTER II

LITERATURE REVIEW

Literature review of the aviation demand and network modeling is documented in this chapter. Research issues are introduced after each subsection, identifying the need of the research.

2.1 Aviation demand forecasting

There are two types of demand estimation models depending on the considered modes of transportation. Multimodal demand estimation tools consider different types of competing transportation modes with the aviation demand modeled as the result of modal split. Single modal airline-only demand models only consider aviation demand without considering competition. These models can also be categorized as models with unknown structure due to their proprietary nature and as the ones with known structure. Some of the most famous aviation demand forecasting models are reviewed in this section.

2.1.1 Aviation demand forecast models with unknown structure

Different organizations have developed their own proprietary demand forecast models. The details of these proprietary models are not well known. But some of them holds authoritative status and their numbers are widely used in aviation researches. The most well known among them are the forecasts made by Boeing, Airbus, and Federal Aviation Association (FAA).

The Boeing Company annually publishes "Current Market Outlook" which contains information about economic issues and air travel growth for world travel (The Boeing Company, 2008). It projects air travel demand and the fleets of aircrafts

required to meet the demand. It has separate projection by international regions, represented by North America, Latin America, Europe, Africa, Middle East, Northeast Asia, Southeast Asia, Southwest Asia, China, and Oceania. It provides gross level information about travel demand and is not appropriate for obtaining information on demand or operations at specific airports in the CONUS. The most recent one is "Current Market Outlook 2008-2027", which shows how air transport will change over the next 20 years.

Airbus also has an equivalent 20 year forecast named "Global Market Forecast" for the international air travel demand and the fleets of aircraft needed to meet the anticipated demand (Airbus S.A.S., 2007). It also provides gross level information about travel demand and is not appropriate for obtaining information on demand or operations at specific airports in the CONUS. The most recent one at the moment is "Global Market Forecast 2007 — 2026".

FAA has a couple of demand forecast models, among which the Terminal Area Forecast (TAF) system is the official forecast of aviation activity at FAA facilities. The TAF provides forecasts for active airports in the National Plan of Integrated Airport System (NPIAS). The TAF includes forecasts for FAA towered airports, federally contracted towered airports, nonfederal towered airports, and non-towered airports. Detailed forecasts are prepared for large air carriers, air taxi/commuters, general aviation, and military. An internet server of FAA provides the historical data and forecasts, which can be queried using any web browser (Federal Aviation Administration, 2009*b*). FAA updates TAF annually and constantly improves the forecast method. The TAF provides forecasts of the following flight activities.

- Enplanements for air carrier and commuter
- Aircraft operations including
  - Itinerant operations for air carrier, air taxi/commuter, general aviation,

and military

  - Local operations (takeoff and landing at the same airport) for general aviation and military

- Instrument operations

TAF summary report is also available to the general public (Federal Aviation Administration, 2008*b*). More information on TAF database is provided in Appendix D.

2.1.2 Aviation demand forecast models with known structure

Aviation demand models in this category have publicly available publications on their processes and structures. In this section, three representative demand-centric models with known structure are overviewed , namely the AvDemand model, the Transportation Systems Analysis Model (TSAM), and the *Mi* model.

AvDemand is a software tool that calculates future NAS demand based on FAA forecasts. It was developed by Sensis Corporation for use by National Aeronautics and Space Administration (NASA) to generate future air transportation demand in Air Traffic Management (ATM) experiments (Huang et al., 2004). It has spatially-explicit representation of the CONUS. The outputs have a generic format so that it can be integrated with other NASA simulation tools. It does not include intermodal and multimodal relationships and it does not capture consumer behavior. AvDemand is composed of three main functions, namely demand generation, demand data input/output, and demand analysis. Demand generation is the core component of AvDemand and offers two approaches: a flight-based demand growth approach and a passenger-weighted demand growth approach. Figure 8 shows the demand generation features of AvDemand. In the flight-based demand growth approach, AvDemand starts with the baseline flight demand set and assumes a growth rate to generate target future demand sets. In the passenger-weighted demand generation approach, AvDemand also starts with the baseline flight demand set and estimates passenger

demand from the flight demand. It then applies passenger growth rates to calculate passenger demand for the target future demand. Flight schedules and flight plans are calculated using a fleet-mix determination and departure time distribution algorithm.

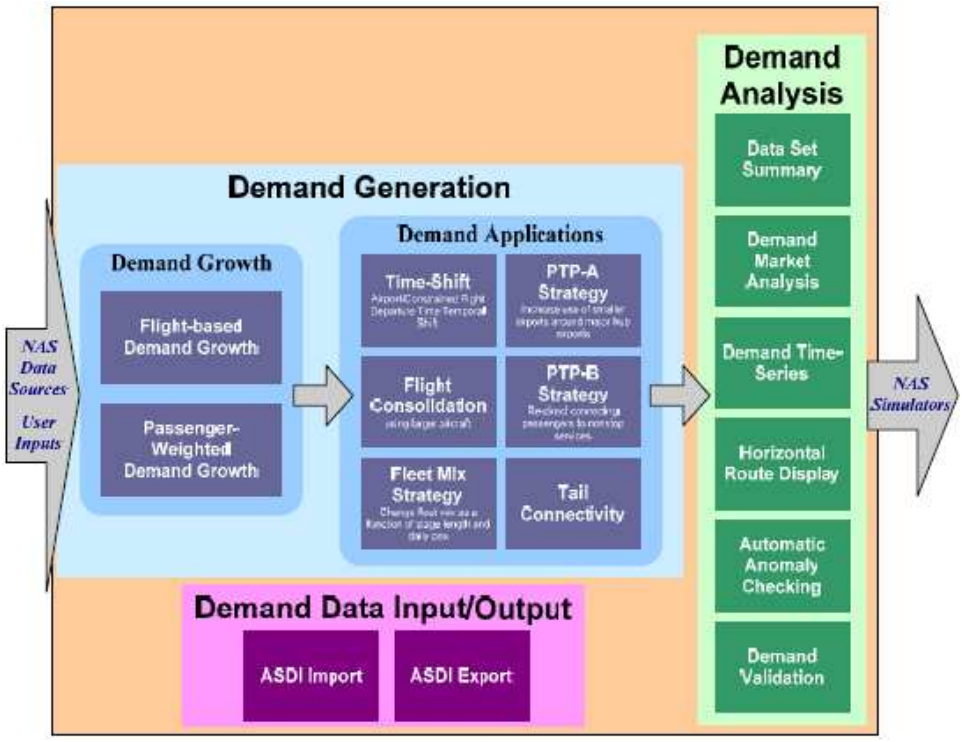


Figure 8: AvDemand demand generation features. [Source: Huang and Schleicher (2007)]

TSAM is a database-driven demand prediction model to estimate long distance travel at a county level based upon population and demographics in multimodal scope (Viken et al., 2006*a*; Trani et al., 2003; Baik and Trani, 2005). It is initially developed by Virginia Tech to study Small Aircraft Transportation System (SATS) program. The mode choice is performed based on the trip purpose, trip cost, time and time value of the trip. TSAM adops a nested multinomial logit model for mode choice. The county level demand for air travel after the mode choice is then aggregated to the airport level. The airport level O-D demand is then assigned to specific routes following the traditional four-step transportation planning framework (More information

on the traditional four-step process can be found in Appendix. A). Figure 9 shows the model framework of TSAM. When future flight demand growth is modeled, TSAM utilizes baseline flight schedules and uses the FRATAR algorithm to develop the future schedules in the NAS (More information on FRATAR algorithm can be found in Appendix. A.3.3). These projected flights can be plugged into air transportation simulators to analyze the impact of the projected demand on the NAS. Viken et al. (2006*a*) used the Airspace Concepts Evaluation System (ACES), which was developed under NASA's Virtual Airspace Modeling and Simulation (VAMS) project (Couluris et al., 2003; Zelinski, 2005; National Aeronautics and Space Administration, 2004), as the airspace simulation program for his research. A description of some of the widely known airspace simulation programs is provided in Appendix B

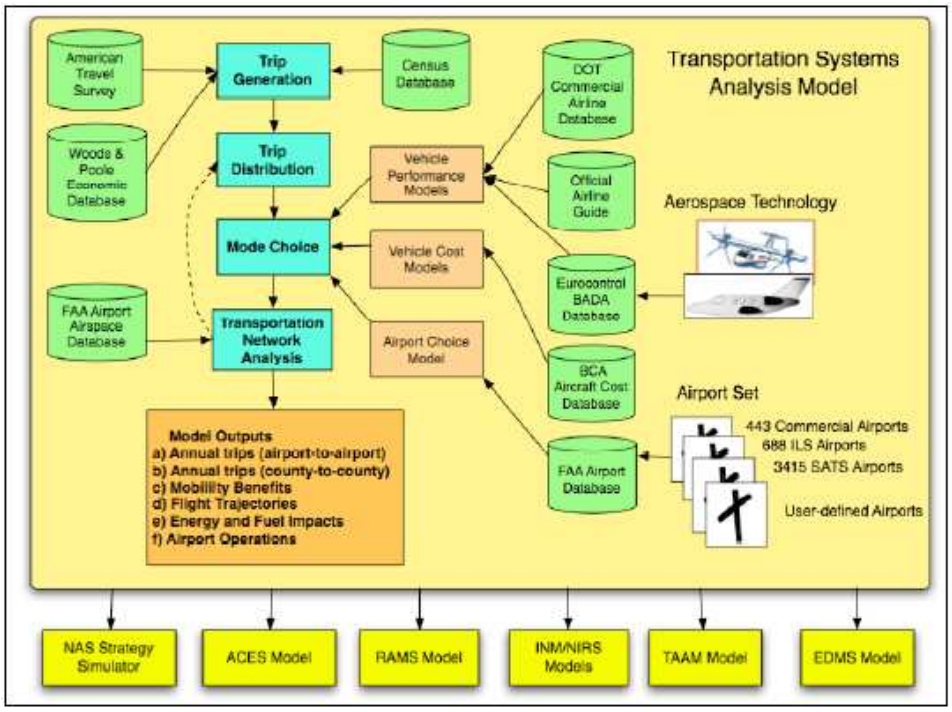


Figure 9: TSAM model structure. [Source: Viken et al. (2006*a*)]

Mi is an agent-based model developed by Lewe at Georgia Tech (Lewe, 2005; Lewe et al., 2006). Both TSAM and *Mi* forecast travel demand at the National Transportation System (NTS) level and consider intermodal and multimodal aspects.

However *Mi* is not spatially-explicit but uses virtual NTS concept, where agents reside and perform transportation activities. The CONUS in *Mi* is represented with four locales, which are large-, medium-, small-, and non-metropolitan areas. Using Agent-Based Modeling (ABM), it tries to capture behavioral aspects of travelers. Its entity-centric abstraction makes the simulation less computationally complicated. It models transportation consumers and service providers as agents. Transportation consumers are modeled based on the demographic and economic characteristics of the locales. Transportation service providers generate price and time information for each mode of transportation and business model. Then consumer agents perform transportation mode choice using a multinomial logit model. Figure 10 shows flow charts of the *Mi* model. "TAF" in the graph stands for Transportation Architecture Field while "S/Ps" represents service providers. The result from the simulation is annual-based and calibration was performed against the 1995 ATS data.

2.1.3 Research issues

AvDemand is a single mode demand forecast tool. Both TSAM and *Mi* are multi-modal demand model, where air transport demand is calculated after modal split. The main disadvantage of having single mode tool is not accounting for other modes of transportation and as a result not considering spillage to other modes when socio-economic property changes occur. But considering the complexity and assumptions that are involved with holistic models which lead to inevitable errors, modeling air transportation only can have benefit of having simpler structures and can leads to equivalent accuracy when properly modeled. Also, AvDemand does not use directional true O-D concept that is necessary for this proposed investigation and TSAM and *Mi* use survey data (1995 ATS), which is exposed to data gathering and human errors. However, real data collected by the government institutions on the air transportation can be used to build a tailored demand model required for this research.

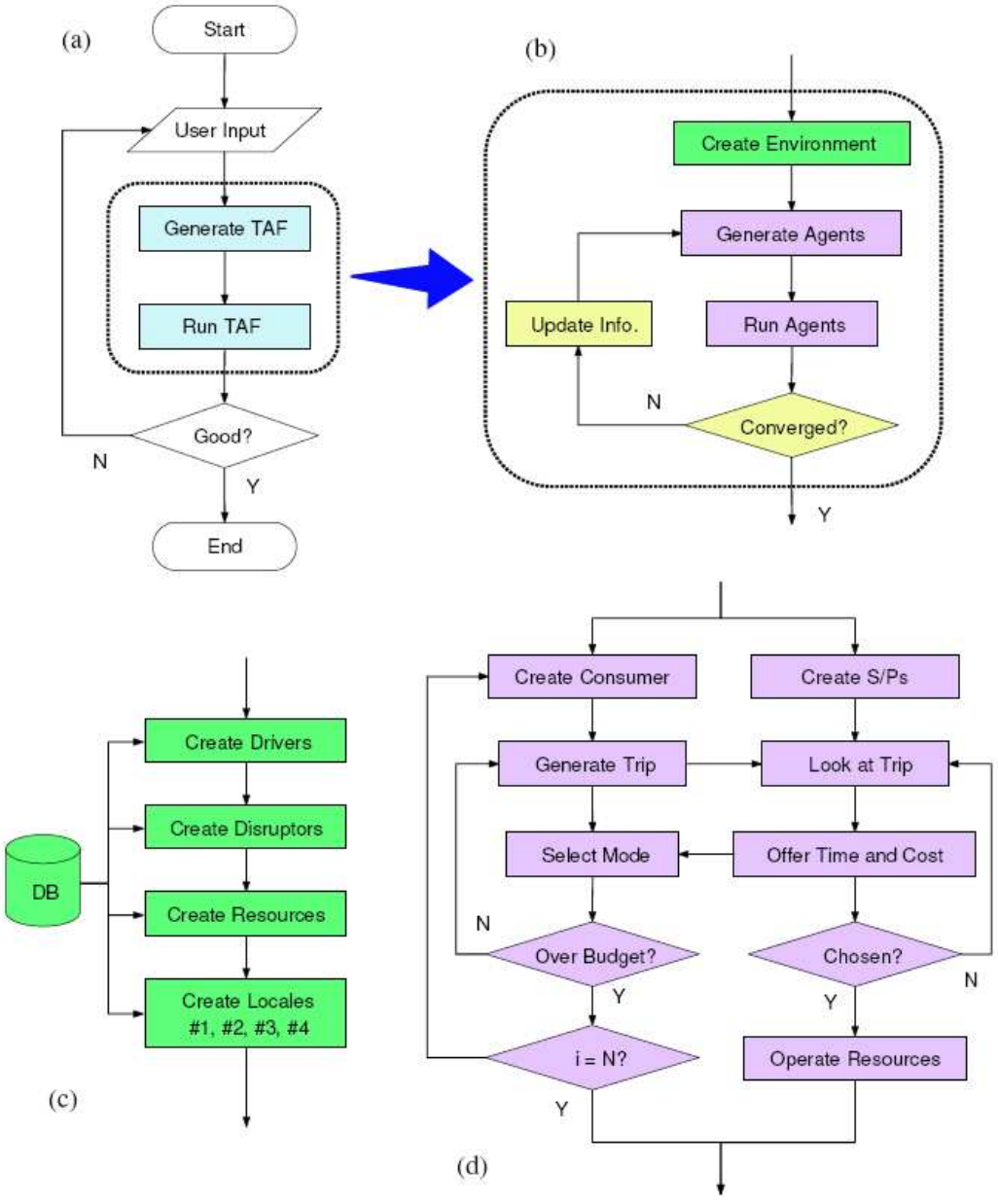


Figure 10: Flow chart of *Mi* model. [Source: Lewe (2005)]

One thing to note is that these data on NAS are also prone to human errors. Some of the erroneous entries found during this research are reported in Appendix D.

2.2 Modeling of complex networks

Complex network is a network with non-trivial topological features as opposed to simple networks such as lattice or random network. Network theory has evolved as a part of graph theory in applied mathematics to explain the fundamentals of the network formation and physics behind it. The study of complex network is a relatively new field and is widely gaining interest in a variety of areas, such as statistical mechanics of complex networks (Albert and Barabsi, 2002), large-scale organization of metabolic networks (Jeong et al., 2000; Ravasz et al., 2002), characterizing the shape of internet (Albert et al., 1999; Faloutsos et al., 1999), network biology (Barabási and Oltvai, 2004), engineering problem-solving networks (Braha and Bar-Yam, 2004), soccer players network (Onody and de Castro, 2004), and epidemic dynamics (Pastor-Satorras and Vespignani, 2001) as well as airline network modeling (Strogatz, 2001; Barabási, 2004). As a result, numerous measures have been developed to characterize networks, including Gini index (Reynolds-Feighan, 2004), degree centrality, betweenness (Freeman, 1978), closeness, influence measure, etc (Borgatti, 2005). Newman (2003) provides a good review on the structure and function of complex networks.

In this section, a comparison is made between evolution-based complex network modeling and airline network analysis to ascertain the benefits of incorporating evolution-based approach on airline network modelings.

2.2.1 Network model with evolution

There has been rapid progress in the statistical physics of evolving networks resulting in many publications that attempt to explain the underlying mechanisms for growing networks (Krapivsky et al., 2000). This research field first started with the consideration of random networks, which has since found much usages in various areas.

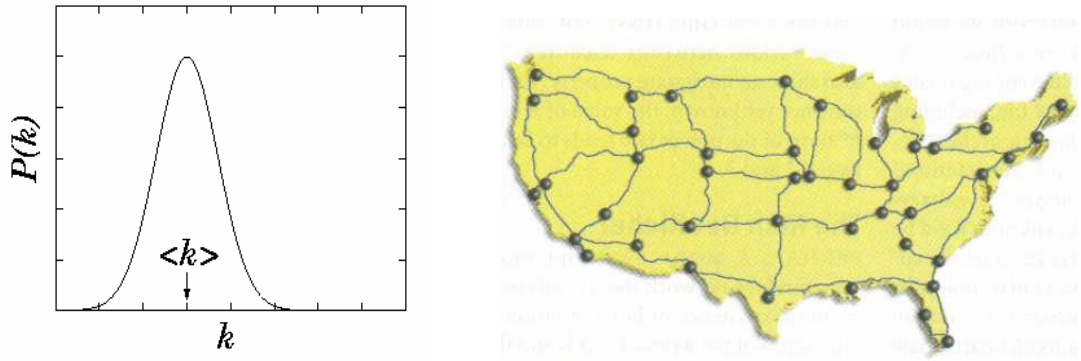
Dorogovtsev and Mendes (2002) published a good review paper on this subject. Two of the most well-known complex networks are scale-free networks (Caldarelli, 2007; Caldarelli et al., 2002) and small-world networks (Cohen and Havlin, 2003). Scale-free networks are characterized by power-law degree distributions while small-world networks are described by short path length and high clustering. In this section, general characteristics of random, scale-free, and small-world networks are reviewed. The possible underlying mechanism for scale-free networks are also introduced.

2.2.1.1 Random network

Random network was systematically studied first by Erdös and Rényi in the 1950s and 1960s (Erdös and Rényi, 1959, 1960; Ravindran, 2007). They proposed uniform random graphs, which is now called ER graph, that does not show specific patterns. It is an evolutionary network that starts with N disconnected nodes. It connects two randomly selected nodes with the probability of an edge between them being $0 < p < 1$. The ER graph is written as $G_{n,p}$, where n is the number of nodes in the graph and p is the probability of connection between any pair of nodes, and is independent of the existence of other connections in the graph. The degree of node 'A' (kA) is defined as the number of connections of 'A' and it is also called neighbors. The average degree of ER graph is found to be $z = \frac{p}{n-1} \approx \frac{p}{n}$ when n is large enough. The probability of a node having degree k becomes $p_k =_n C_k \cdot p^k \cdot (1-p)^{n-k} = \frac{e^{-z} \cdot z^k}{k!}$. There are other random graph models developed by other researchers. For example, the $G_{n,m}$ model describes a graph that is composed of n nodes with m randomly selected edges. Different models generate different degree distributions but the connections of random networks are completely random, implying that the degrees of all the vertices are integers drawn independently identically from specific distributions. After Erdös and Rényi, a good number of interesting publications in the evolutionary network research were proposed.

There are physical systems that demonstrate abrupt phase changes at certain points such as liquid/gas, magnetization, superconductivity, etc. One of the most interesting phenomena of a random graph is that it also shows a distinct phase transition. A Giant Component (GC) is defined as a connected subgraph that contains a majority of all the nodes in the graph. If the network evolves with increasing p, after a certain point in the evolution, the graph experiences phase transition from a disconnected to a connected network with emergence of GC. From theory, it is expected that most of the connected components are small with the largest having $O(log(n))$ vertices when $p < 1/n$. Sharp transition occurs at $p = 1/n$ and the largest component contains a finite fraction (F) of the total number of vertices, $C_{max} = F \cdot n$, and all other components remain at the size of $O(log(n))$. When p is close to $log(n)/n$, the graph is expected to be completely connected.

Random networks have a single hump when its connectivity is represented with distributions. The nodes at the peak of the distribution are called the characteristic nodes and make up the majority of nodes with the same connectivity in the network. Only a small number of nodes lie outside this hump range. Random network is also called an exponential network since the degree decays rapidly, i.e. exponentially. The most representative exponential network in the U.S. transportation system is the highway network. Highway network resembles grid network that is also characterized by the existence of the characteristic nodes. Figure 11 shows the characteristics of the U.S. highway system.

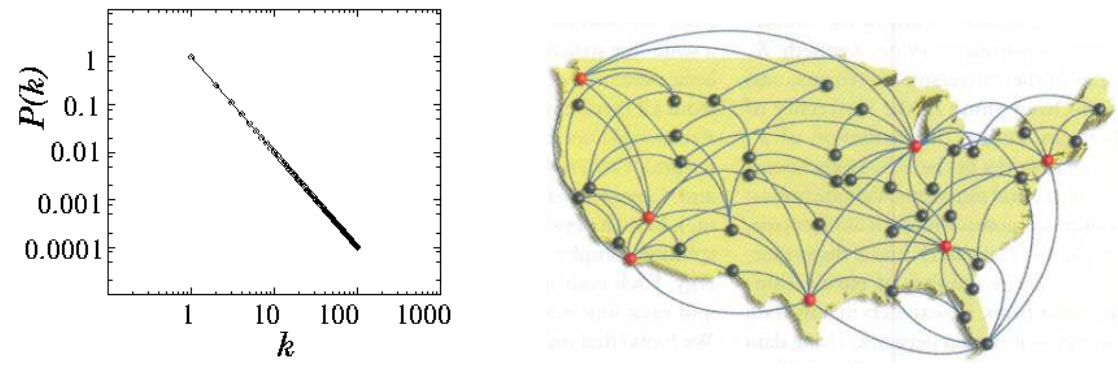


(a) Poisson distribution (b) U.S. highway network

Figure 11: Random network. [Source: Barabási and Bonabeau (2003)]

2.2.1.2 Scale-free network

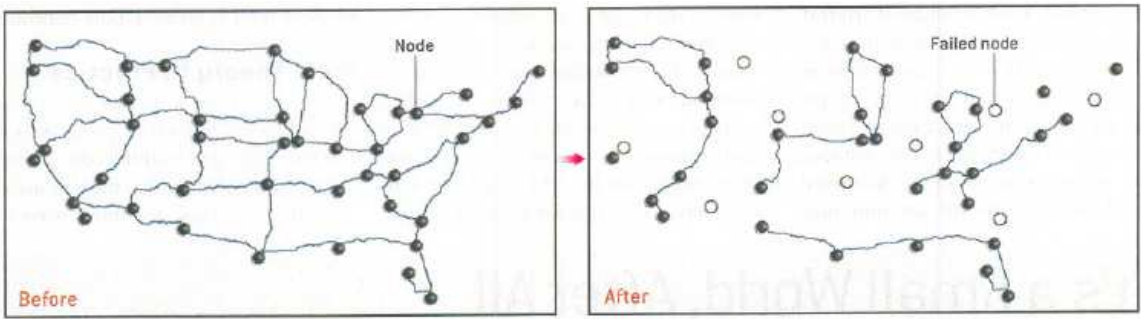
Unlike random network, most network systems that are observed in nature do not exhibit purely random behavior. Instead, the majority of them show power-law behavior of the degree distribution (heavy tail) and clustering. If the degree distribution follows a power-law, it does not have a peak in the distribution. Power-law distribution is described as most nodes having few connections and some nodes having a lot of connections. In other words, all levels of degree size exists or also referred to as "no" scale (or scale-free). The nodes with very high degree in a scale-free network are often called 'hubs'. There are many processes that generates a power-law degree distribution. Figure 12 shows an example of scale-free network.



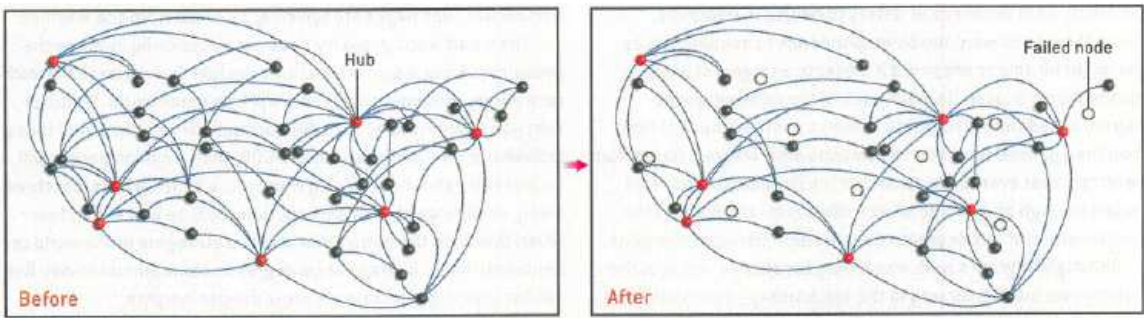
(a) Power-law distribution (b) Scale-free Network

Figure 12: Scale-free network. [Source: Barabási and Bonabeau (2003)]

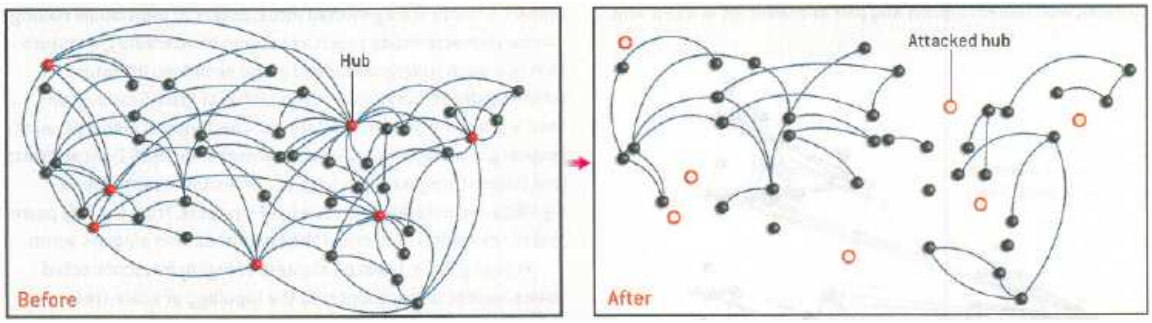
Scale-free network exhibits robustness under random disruptions. However, when malicious attack is inflicted on the hubs, scale-free network exhibits catastrophic effects due to the concentration of network flows (Albert et al., 2000; Tu, 2000; Jeong et al., 2001). Figure 13 shows an example of random failures and malicious attacks on the hubs. The vertices with high degrees are a special concern because they have critical roles, such as in the spread of diseases, the transmission of packets in the internet, or the bottleneck problem in the air transportation system.



(a) Random network - Accidental node failure



(b) Scale-free network - Accidental node failure



(c) Scale-free network - Attacks on hubs

Figure 13: Failure: random vs. scale-free network. [Source: Barabási and Bonabeau (2003)]

Scale-free networks are frequently observed in real life, including world wide web, social, protein, and citation networks. The most representative scale-free network in U.S. transportation system is airline network. Many researches have proposed underlying mechanisms for scale-free networks (Newman, 2004; Barabási et al., 1999). The idea of preferential attachment is a "rich-get-richer" philosophy, meaning vertices with more connections gets a higher chance of being connected. Preferential attachments has been proposed as an underlying mechanism of power law distribution of connectivity in networks (Dorogovtsev et al., 2000), which include the distribution of the city sizes (Simon, 1955), the citation network (Price, 1976), individual wealth distribution (Price, 1976), and the number of pages linked to certain pages on the world wide web (Barabási and Albert, 1999; Huberman and Adamic, 1999). Figure 14 shows a model proposed by Barabási and Bonabeau (2003) that generates scale-free network by growth and preferential attachment. In this model, a node is added each step (growth) and the introduced node is attached to two existing nodes with the linear probability of connection (preferential attachment), linear meaning that a new node is twice more likely to be attached to an existing node if it has twice as many links as others. Barabási (2004) claims that this is the first model to explain the scale-free power laws seen in real networks.

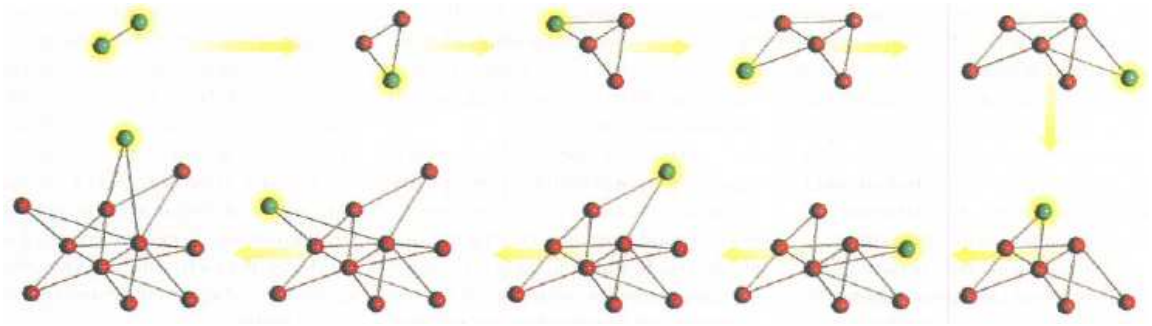


Figure 14: Growth and preferential attachment generating scale-free network. [Source: Barabási and Bonabeau (2003)]

Since it is a fairly new research field, there are ambiguities in the terms used by

different researchers. In order to clarify the definition of scale-free networks, Li et al. (2005) published a technical report providing precise definitions.

2.2.1.3 Small-world network

Small world network refers to a network where arbitrary two vertices are separated by small number of connections. The name comes from an analogy with the small-world phenomenon, which was first proposed by Frigyes Karinthy in 1929 and tested by Stanley Milgram in 1967. It describes any two people on a social network are only separated by six degrees of separation. In other words, the diameter of the network is six. Watts and Strogatz (1998) published the first small-world network model. A wide variety of networks including random and scale-free network show characteristics of small-world network.

2.2.1.4 Network value

When there is a network with N nodes, $\frac{1}{2}N(N-1)$ possible interconnections exist. If one connection has the same value as the others, the total network value is $K \times N(N-1)$, where K is a constant. So, with the addition of one user to the network, the total network value increases at a rate approximately proportional to N^2 when N is large, resulting in the fact that the value of the network per node being $N-1$ rather than a constant. This is often referred to as Metcalfe's law which means that one additional node gives bigger value than one to the network.

Metcalfe's law implies a "critical mass" phenomenon. If number of users is smaller than the "critical mass", the network usage decreases until it finally dies out, while with a greater number of users, the usage grows. This observation applies to the Internet development. After some time from the inception of Internet, its usage reached critical mass. As the World Wide Web and Web browsers evolve, it attracted more users and content providers added more services to the Internet. As the online communities grew — like online stores, government services, and activity clubs — the

value of the network increased and the Internet continued attracting more and more users, resulting in again, more value for each user, and so on. Metcalfe's observation is also described as the "network effect". Network effect is a phenomenon where the value of a product or a service to the user is an increasing function of the number of users. In transportation systems, there exist direct and indirect network effects. A direct network effect is a case when there is direct relationship between the usage and the value of the network. An indirect effect plays behind the scene. A network effect in the air transportation network can be expressed using a virtuous cycle as in Figure 15, with situations that have positive effect on one another. Figure 15 includes both direct and indirect network effects.

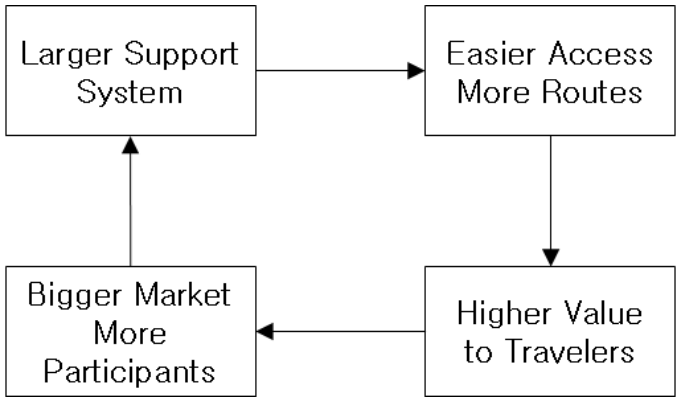


Figure 15: Air transportation network effects virtuous cycle

A direct effect shows that increasing the number of connected cities connected through the new transportation service also increases the value of the network. An indirect effect is not as obvious but can still be noticed. For example, the usage of the new system not only depends on the number of connected cities, but also on the convenience and mobility of the transportation system. Travelers choose a service that brings the highest value to him or her from among the available options. As the new system gains more support, it will bring a higher value in terms of convenience and mobility to the traveler making it more likely for the traveler to find the best value with the service. This positive feedback loop induces more of the system adoption

and support in the community until the new system dominates the market or the equilibrium point among the services is reached. So, increased installed base induces an increase in number of support systems available, and the resulting increased service induces more installed base.

2.2.2 Airline network modeling

Airline network modeling takes a form of flight scheduling (Brueckner, 2004; Brueckner and Zhang, 2001) and fleet routing (Dobson and Lederer, 1993; Desaulniers et al., 1997; González, 2006) problem and is generally set up as a cost minimization (Yan and Tseng, 2002; Adler and Berechman, 2001) or profit maximizing problem (Lederer and Nambimadom, 1998). In this respect, Weidner (1996) developed a model that explains airline hubbing in the NAS. Yang and Kornfeld (2003) formulated package delivery system also in the context of cost minimization to show how H&S and P2P structures are formed under various environments. Hsiao and Hansen (2005) developed an equilibrium model for air transportation network flow. Hsu and Wen (2000) applied Grey theory and multi-objective programming method to develop a model capable of designing airline network. Adler (2005) developed a model that analyzes the H&S network design issue under competition in a game theory framework. There are numerous studies and simulation efforts on the NAS (National Aeronautics and Space Administration, 2004; Niedringhaus, 2004). There are also numerous other studies that describes airlines' choice in network structure (Oum et al., 1995), with similar mathematical settings to solve for equilibrium network. As a tool to characterize the network, many network measures have also been developed (Alderighi et al., 2007). There have also been attempts to address the relationships between demand uncertainty and airline network structure (Barla and Constantatos, 2000). As evidenced in the literature, airline network modeling and analysis have been done with equilibrium network, not evolutionary ones.

2.2.3 Research issues

It was identified that valuable insights from graph theory has not been extensively exploited in air transportation network models (Kotegawa and DeLaurentis, 2007). The need for airline network modeling with a graph theory approach was also discussed in a workshop on transportation network topologies organized by the NASA Langley research center (Alexandrov, 2004). Researchers promote preferential attachment as possible explanation for power-law distribution in many network systems and model them with evolutionary nature, instead of modeling networks for equilibrium, which is popular among researchers in the transportation field (Marcotte and Nguyen, 1998). Even though airline network clearly exhibits power-law distribution, practices of airline network modeling have not included the evolution scheme. Therefore, the possibility of enhancing the state-of-the-art modeling technique using evolution scheme needs to be investigated. However, DeLaurentis et al. (2008) pointed out that few airports entered the NAS and the rerouting process dominated the evolution in the recent years, making Barabási-Albert (also called BA) model not appropriate for the airline network modeling. Hence, different evolution scheme should be needed in the context of airline networks.

2.3 Literature review summary

From the reviewed data, theories and models, it can be concluded that the current data and models are insufficient on their own to realize the motivation of building an active design tool because the demand models do not have true O-D information and the network models do not incorporate evolutionary growth. Additionally, the NAS evolved without much control similar to the world wide web. Researchers have shown that there is order hidden in the web: their connections follow a universal power-law. Similarly, the NAS also have universal power-law connections and may be explained

using a preferential attachment in a growing network model. Therefore, a new approach needs to be formulated. In this thesis, an active control algorithm for the air transportation network is formulated by building a demand model that utilizes true O-D concept and a network model that grows with preferential attachment scheme.

Accuracy is always a big concern when building an engineering model. This is particularly so for air transportation system as the cost involved with infrastructure construction is high. However, a study by Flyvbjerg et al. (2006) provids with high statistical significance that demand forecasts generally have been done poorly for infrastructure planning projects, even for highly detailed simulations. For example, the TRANSIMS project by Los Alamos National Laboratory did not show much benefits in accuracy. On the other hand, the researches on complex network modeling show that simple and elegant characterization can open possibilities to the analysis of complex system. Therefore, this research will try to capture the essence of the true nature of demand and network properties in an airline network with the simplest possible setting.

2.4 Research questions and hypotheses

The research objectives were introduced in Chapter 1, while the current practices and theories were reviewed in this chapter. Research questions and hypotheses that were raised during the course of the research are presented.

2.4.1 Research Question 1: How can the intrinsic demand model be generated?

The real NAS cannot be known perfectly because of its size and complexity. However, it is reasonable to think that its aspects are captured in the data collected by BTS. The two public data that can be used for building a true O-D based demand model is the DB1B and T100. DB1B is a 10% sample that tracks individual trips, while T100 is an aggregated data reported by all the certified airlines operating in the U.S..

Therefore, T100 captures the overall picture of the NAS but does not include detailed trip information. Hence, it is proposed to build a NAS demand model based on socioeconomic data in order to generate the trends captured in the detailed sample data (DB1B). The result is then compared to T100 to ensure that the small-scale model is representative of the full-scale meta world. So, the hypothesis for the first question is as follows.

Hypothesis 1:
A demand prediction model that predicts intrinsic need can be built based on socioeconomic characteristics.

Having dealt with the demand model, the next research question is as follows.

2.4.2 Research Question 2: Is airline network evolutionary?

This question is necessary to select the best structure for an airline network model. It is widely agreed that the cost associated with changing the topological usage of the airline network is very high. When there is resistance, the previous state has a great amount of influence on the next state (in other words, current state is input to the next state), suggesting a stepwise evolution scheme. Many evidence also points to the fact that the airline network is an evolutionary network (See Section 3.1). Since many evolutionary networks in nature are explained by the preferential attachment scheme, it is reasonable to apply preferential attachment as a underlying mechanism for the proposed evolutionary network model. Therefore, the following hypothesis is made.

Hypothesis 2:
Airline network is evolutionary and can be explained with preferential attachment.

The next sub-question arises naturally about how to implement preferential attachment in airline network modeling.

2.4.2.1 Research Question 2-1: How can preferential attachment be implemented in airline network?

This question raises attention on the difference of airline network and abstract networks in graph theory. Airline network is a physical network formed by the operation of aircrafts between airports. When economies of scale are realized by using bigger airplanes or there is available capacity at certain segments, demand on the segments are fortified. Therefore, airline network modeling is better approached differently from general preferential attachment scheme typically used in abstract network modeling. In this research, a network is built by introducing the most demanded airports first and evolving them to a certain maturity, before introducing the rest within a very short time period in the evolution timeline. This idea can be explained by critical mass in Matcalfe's law reviewed and supported by empirical observation of most evolutionary networks (Mendelson, 2003). The following hypothesis is formed to realize preferential attachment scheme in airline network modeling.

Hypothesis 2-1:

Airline network can be evolved by adding highly demanded vertices first and evolving them to a certain maturity before adding the rest within a very short timeline.

CHAPTER III

APPROACH

Based on the motivation (Chapter 1) and findings from the literature review (Chapter 2), some research questions and hypotheses were made (Section 2.4). In this section, a formulation of the aviation demand model and the active design algorithm are presented. As with most cases, models are applicable only under expectable circumstances. Radical or catastrophic events are often not captured in these general models. Similarly, this formulation is limited to the normal case and should not be applied to anomalies in the market. Therefore, validation of this formulation can be applied to the data until $September 11^{th}, 2001$ and significantly after it.

The focus of this research is to build an evolutionary airline network model through preferential attachment. The three major components of air transportation network capacity are the airlines, the air traffic control, and the airports (Merlis, 2001). It is also assumed that there is appropriate air traffic control such that the impact from air traffic control is minimal. This is a reasonable assumption as traffic control related delays are not common in the United States. Therefore, only the airlines and airports are modeled in this research.

Domain analysis is a crucial element that leads to the abstraction of a system. Due to the different ways of representing a selected system, the resulting abstractions do not normally converge into a unique model, especially for a complex system. Since many appropriate models for a select system can exist, there is no definite way to conceptualize a system. In light of that, the approach taken in this thesis is not the best but that it represents a novel approach in the modeling of evolutionary airline network. This is achieved through an active design algorithm that expands the design

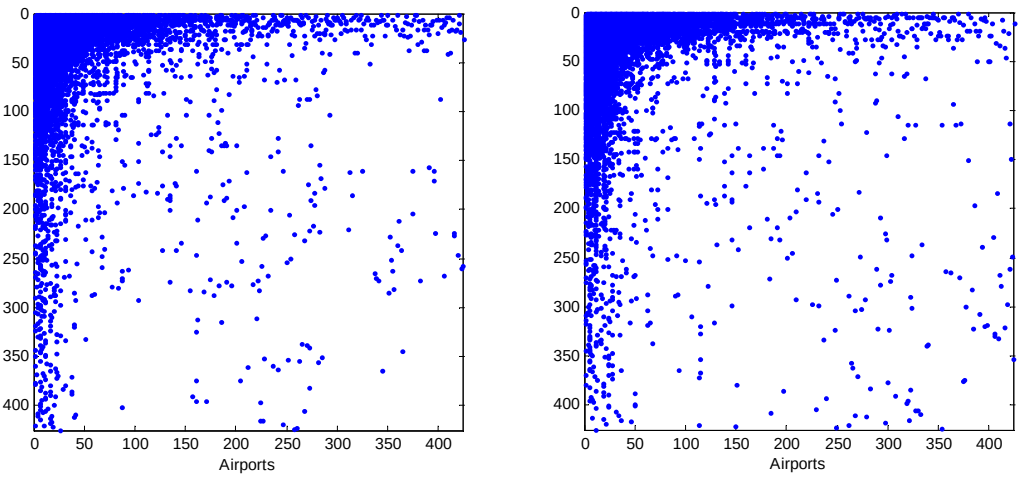
space of the NAS.

3.1 Investigation of the NAS

An investigation of the NAS characteristics was performed based on the available data.

3.1.1 NAS decomposition by connectivity

The NAS is characterized by tightly-coupled interactions between consumer and service provider in the transportation system. These complex interactions make it difficult to fully observe and understand the underlying behavior of the transportation network. To investigate the network-level characteristics Figure 16 was drawn for the NAS in 2005. Figure 16a shows the connection matrix ordered by the number of connections and Figure 16b shows the operations matrix ordered by the number of operations, highest being located on the top left corner. It can be observed in the figure that the NAS is a highly concentrated H&S network and ordering by connections or operations results in almost same graph.



(a) Ordered by number of connections (b) Ordered by number of operations

Figure 16: NAS map in 2005. Points on the plot indicate airports as a connected origin (horizontal) - destination (vertical) pair

To find out if the same characteristics are observed from different perspectives,

a decomposition of the NAS by cliques was attempted as shown in Figure 17. The largest clique that is fully-connected to one another (maximal clique) is categorized as the top level (α-level). The remaining airports are classified with respect to their connectivity to the α-level airports. β-level airports are airports that have connections to any of the α-level airports and ω-level airports are the ones that do not have any connection to any defined α-level airports. A sample result of this process for 2005 is shown in Figure 18. Figure 18a shows the map of connections that are ordered by number of operations and it also shows the characteristics of the 'rich-get-richer' network. In studying the evolution of the NAS network, annual changes in the size of each network level were investigated and depicted in Figure 18b. Examination of the connection matrix for each year (not shown) suggested that both H&S (α-β connections) and P2P travel (β-β, β-ω, ω-ω connections) are growing faster than hub-to-hub (α-α connections) travel.

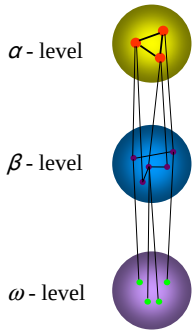
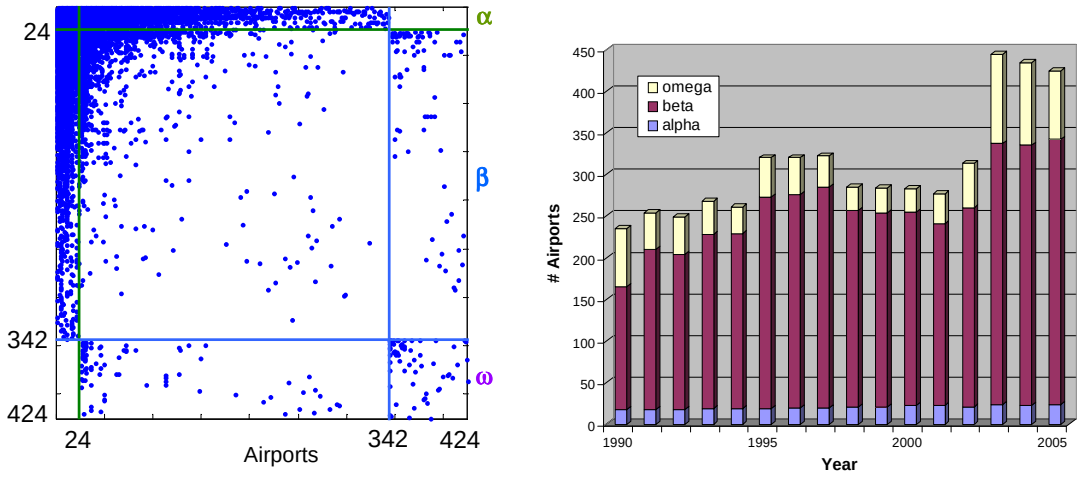


Figure 17: Network decomposition by maximal clique. [Source: Lewe (2008)]

Several assumptions were made in order to generate the α-, β-, and ω-levels networks. First, only airports in the CONUS are considered. Second, the large list of NAS flight segments is truncated by considering only airport pairs with greater than or equal to 365 annual scheduled departures. Third, all the airports are sorted in ascending order based on the total number of outgoing operations.

In an attempt to see the evidence of the 'rich-get-richer' network in local levels, the previously decomposed network using maximal clique was subdivided using



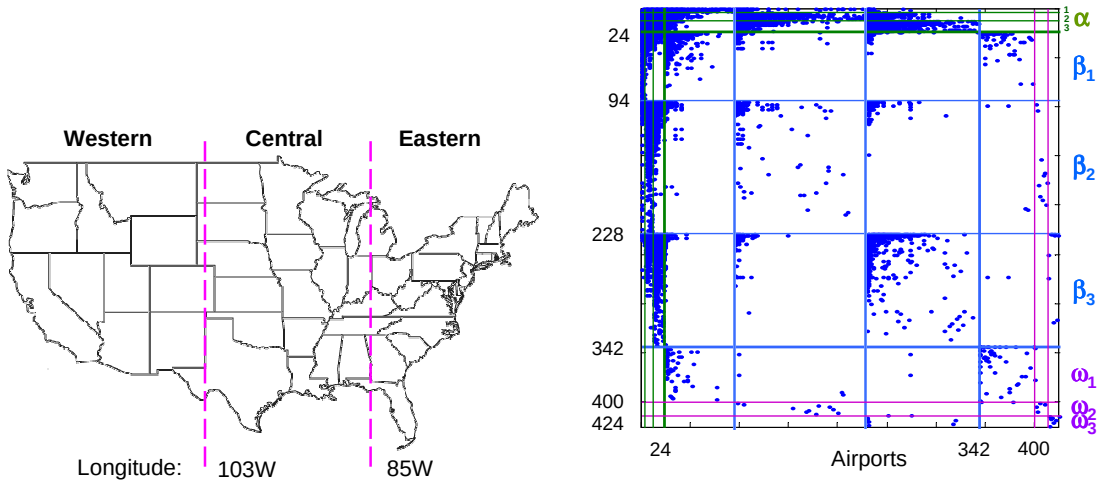
(a) Ordered by operations

(b) Yearly changes of the number of α-, β-, and ω-level airports

Figure 18: NAS network connections, grouped by α-, β-, and ω-level, 2004

geographical split. Figure 19 shows the result of the network decomposition for a subdivided NAS. For this purpose, the CONUS is divided into three regions, namely (1) Western, (2) Central, and (3) Eastern. As expected, it is observed that the connections between airports within the same region are much denser than with other regions either at the same or different network levels. This implies the significant influence of physical landscape and geometric distance in shaping the NAS, particularly for airports represented by the β-level networks, which tend to be smaller hubs and regional airports. These β-level airports are also heavily connected to α-level airports but sparsely connected to ω-level airports, indicating strong preferential attachments existing in the NAS. Since airports at each network level are sorted in the order of operation levels, the distinct gathering of dots towards the upper left corner of each region forms a natural boundary for the region, implying that airports with higher demand have more connections between them.

After visually inspecting the connections between airports in the NAS, the operations and the passenger demand levels were investigated to find more evidence of preferential attachment.

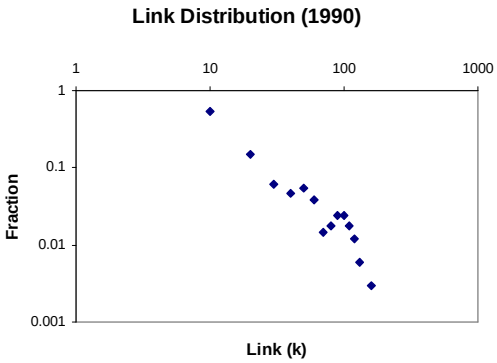


(a) Dividing lines (b) Subdivided plot

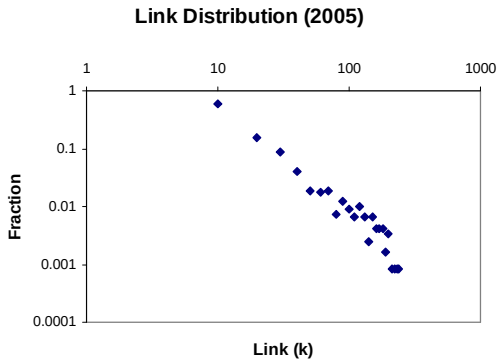
Figure 19: NAS network connections, grouped by α-, β-, and ω-level, subdivided by region, 2004

NAS has evolved into a distinct H&S system. H&S system are best explained by preferential attachment in a growing network. A network that has grown by preferential attachment scheme shows power-law distribution (See Chapter 2) and historical observation of the NAS shows this power-law distribution. Figure 20 shows power-law behaviors of connected links, operations, and enplanements at airports on NAS for 1990 and 2005. The bin sizes for the frequency aggregation in this study were 10 for links, 20,000 for operations, and 1,500,000 for enplanements. Since there are many airports that have very little usage (for example, 1 operation per year), the operations and enplanements plot considers airports with more than 730 operations per year (2 operations per day average) and 7,300 enplanements per year (20 enplanements per day average). Even with the cutoff of 730 departures and 7,300 enplanements, the graph shows the emergence of low-demand airports in 2005 that deviates from general power-law.

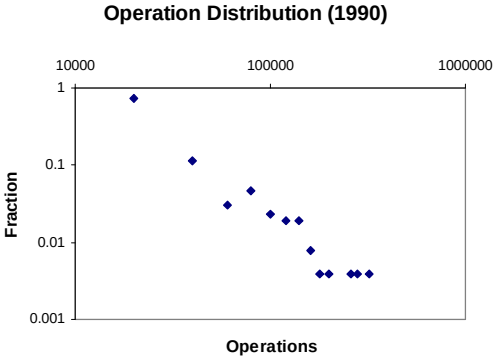
As introduced in Section 2.2, this power-law distribution indicates preferential attachment. A hypothesis that the current NAS can be simulated by preferentially attaching airports in a growing the network is put forth and an active design algorithm



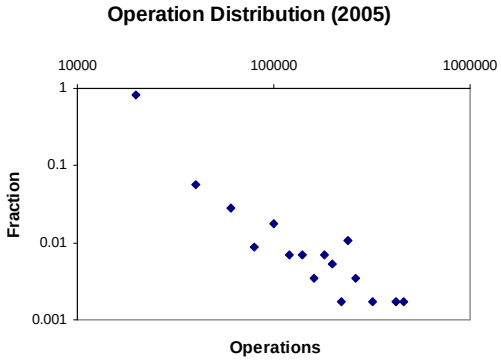
(a) 1990 #Links



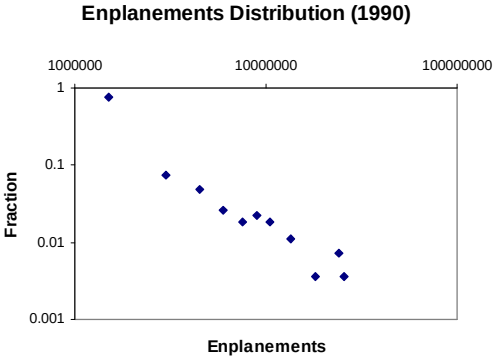
(b) 2005 #Links



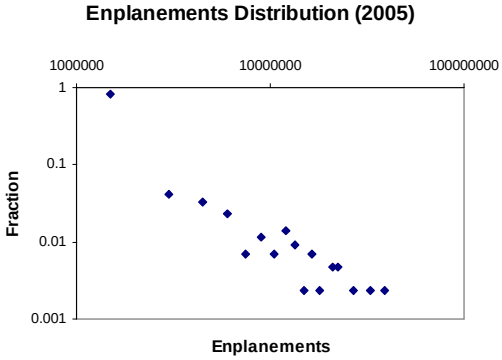
(c) 1990 #Operations



(d) 2005 #Operations



(e) 1990 #Enplanements



(f) 2005 #Enplanements

Figure 20: Power-law behavior of links, operations, and enplanements of the NAS

is created based on it. Also, it was identified from this study that there should be a mechanism in a network model that can give more or less attraction to preferential attachment scheme to accommodate the fact that the NAS has airports that deviates from power-law behavior. Therefore, another sub-hypothesis is added to answer the research question 2-1, which are reorganized with the research question as below.

Research Question 2-1:

How can preferential attachment be implemented in airline network?

Hypothesis 2-1-1:

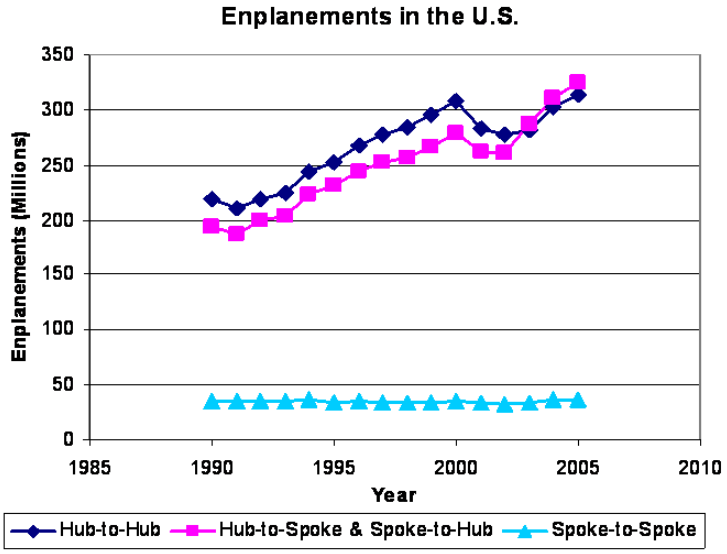
Airline network can be evolved by adding highly demanded vertices first and evolving them to a certain maturity before adding the rest within a very short timeline.

Hypothesis 2-1-2:

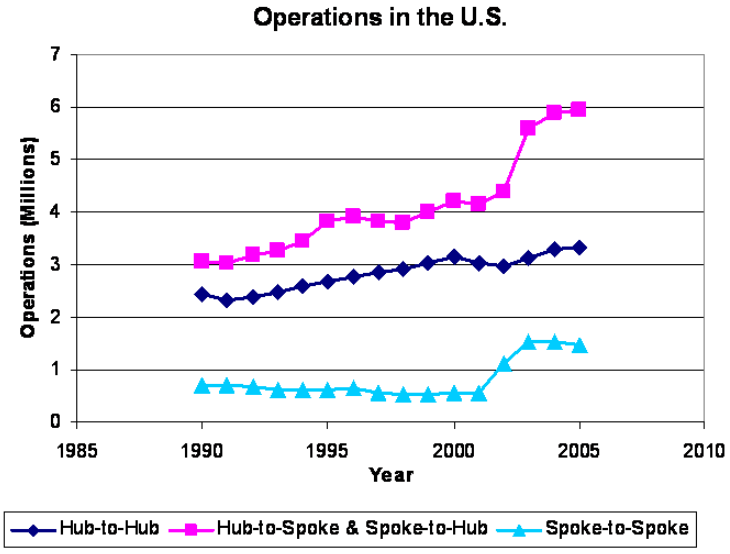
Many airports in the NAS deviates from general power-law. By introducing attraction factors and fine-tuning the attractiveness on a growing network, the real transportation network can be emulated.

3.1.2 Other findings from database study

Figure 21 shows enplanements and operations in the U.S.. Here, airports that are part of OEP 35 airports are categorized as hub airports. The trend on Figure 21 (a) shows that total enplanements have been increasing rapidly in the hub-hub and hub-spoke connections, due to a higher degree of hub concentration. The operations trend in Figure 21 (b) suggests that operations between hub-spoke and spoke-spoke have increased rapidly recently. This results reflects the observation in the airline industry where full-service carriers choose H&S to cover all markets while low-cost carriers select P2P. Both of these services are increasing in quantity. The coexistence of these business models can be explained with the game-theory framework developed by Alderighi et al. (2005). The emergence of low-cost carriers bring many research interest to this field (Morrell, 2005; Reynolds-Feighan, 2001).



(a) Enplanements trends



(b) Operations trends

Figure 21: Enplanements and operations trends in the NAS
* OEP35 airports are denominated as Hubs. All other airports are Spokes.

$Source : T100Dsegmentdata$

Another interesting trend is observed. Figure 22a shows flight O-D pair count of regional, national, and major airlines. Certified air carriers are categorized by BTS by annual operating revenues as in Table 5.

Table 5: U.S. Air Carrier Categorization by Bureau of Transportation Statistics

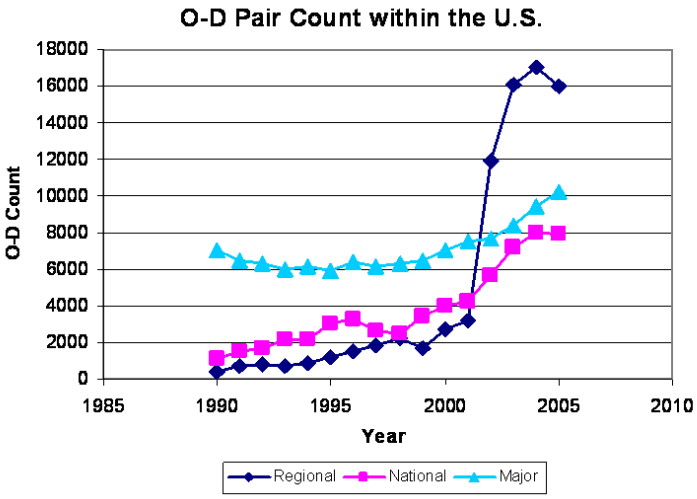
Code	Carrier group	Annual operating revenues
1	Large Regional Carriers	$20 million to $100 million
2	National Carriers	$100 million to $1 billion
3	Major Carriers	Over $1 billion
4	Medium Regional Carriers	Up to $20 million

Figure 22b shows service artery length. (Here, the term 'artery length' is defined as the summation of all the O-D pair distances without duplication.) The postulation for this phenomena is that many previously un-serviced small markets are now serviced by regional airlines.

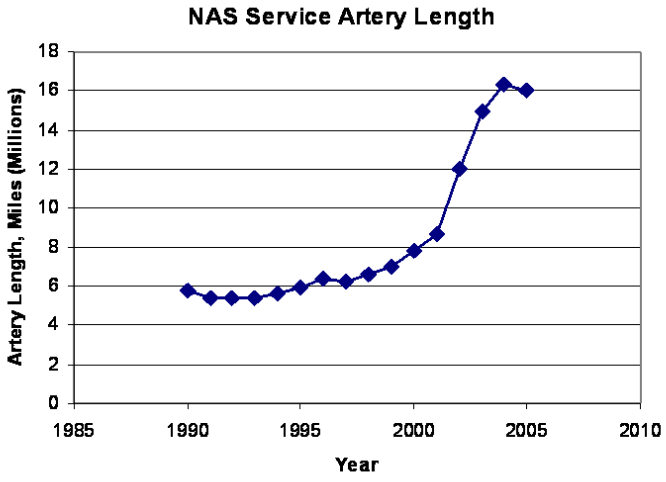
The above phenomena shows the trend of hub concentration and increased hub-spoke network traffic on the NAS. Though true O-D demand is different from enplanements demand and H&S from P2P network, there is a lack of research that deals with the true O-D demand and network effects at the same time.

3.2 Aviation demand model

This section presents an approach that addresses the limitations of current practices on aviation demand. As opposed to segment and airport enplanement metrics used in many studies, 'true' O-D demand is a focal point in this research. An aviation demand model analogous to the gravity law is hypothesized to uncover the underlying mechanics governing produced and attracted enplanements at a local airport. Also, the granularity of the model is on the Metroplex Airport Operation (MAO) level since airports service the base demand of their region together. MAO is a recently emerging concept across the air transportation community aiming to achieve the effective use of adjacent airports where air traffic demand is the greatest. There is



(a) O-D pair counts



(b) Service artery length

Figure 22: O-D pair counts and service artery length. [Source : T100D segment data]

another term that is widely used in the aviation community called Multiple Airport System (MAS), which describes the trend of emergence in the vicinity of primary airports when demand reaches the limit of the primary ones. The emergence of secondary airports leads to the development of multiple airport system. Hansen and Weidner (1995) used 50 km (30 miles) as its boundary and Bonnefoy and Hansman (2007) used 50 miles in their analysis. MAO is different from MAS in that it deals with operational synchronization of airports. The MAO vision is one where runways are dynamically allocated and interconnection between airports are established. However, the technical challenges associated with this new concept of operations are not trivial. Three most significant challenges are (1) restructuring of the NAS and dynamic allocation of existing runways, (2) equalization of security issues between small and large airports, (3) fast ground transportation for the rapid movement of passengers, baggage, and crew. In addition to the technical challenges, the concept creates many other challenges in different dimensions such as political, economical, and societal arenas. Since the implementation requires careful coordination between public and private organizations, impact analysis of the plan needs to be performed to justify the anticipated investment. By building a true O-D based demand model, this research is expected to increase the accuracy of demand prediction of the NAS and help unconventional studies, such as the MAO.

3.2.1 Concept and terminology

Before discussing the aviation demand, related concept and terminology are illustrated in Figure 23. Directional true O-D demand (t_{ij}) represents the number of outgoing trips from origin i to destination j. This information can be collectively presented in the form of a N-by-N trip matrix, where N indicates the number of locales considered. Sum of elements on row i of the true O-D matrix is the produced

enplanements of airport i (P_i) and sum of elements on column j is the attracted enplanements of airport j (A_j). In most cases, t_{ij} is not the same as t_{ji}. However, the O-D demand that needs to be served in both directions are the same since travelers eventually return to their origin location. It is this symmetric demand τ_{ij} ($= t_{ij} + t_{ji}$) that the airlines have to serve, thus becoming the focus of the aviation demand study. Even though the O-D demand (τ_{ij}) is the demand to be served, the true O-D demand (t_{ij}) should be the basis of demand measure since t_{ij} is not tractable from τ_{ij} unless additional information on either t_{ij} or t_{ji} is available (It should be noted that the true O-D matrix (t_{ij}) is different from O-D matrix (τ_{ij}) or enplanements matrix (E_{ij})).

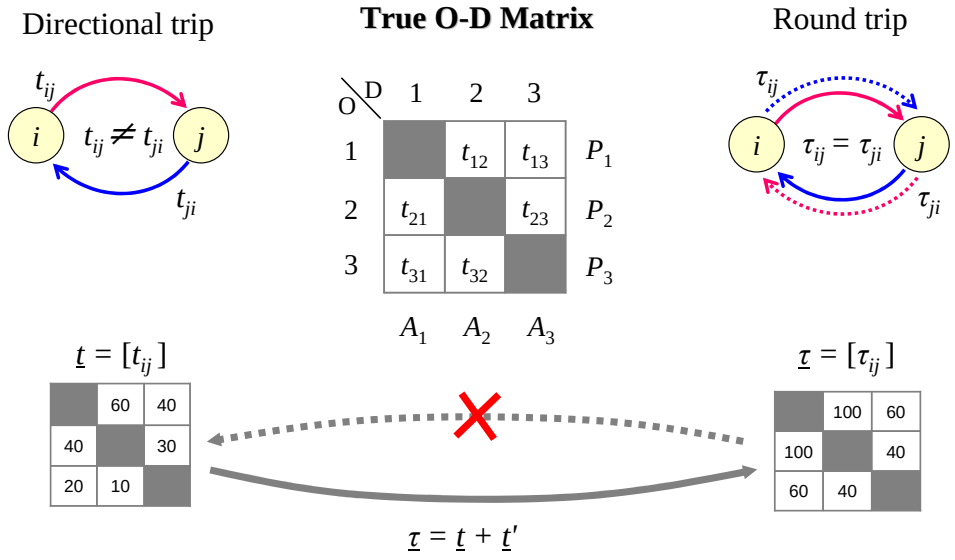


Figure 23: Trip demand terminology. [Source: Lewe (2008)]

The H&S network formed by airlines as a result of economic concerns makes the airline network unique. For the case of perfect P2P service, the O-D demand (τ_{ij}) becomes identical to the enplanements demand (E_{ij}) at each segment. On the other hand, if some of the travelers have connections, the O-D demand and enplanements become different and the total sum of the enplanements on the network becomes higher than the total sum of the O-D demand. To aide in demand analysis, a characterization scheme called the 'PACE' breakdown is introduced as illustrated

in Figure 24. In this terminology, enplanements at airport i (E_i) are divided into produced (P_i), attracted (A_i), and connecting enplanements (C_i) at i. The symbol representation used in the figure are:

- Trip starts from an open end
- Connected lines represent one itinerary
- Filled circles indicate origin (O) or destination (D) where a trip break occurs
- Hollow circles indicate a connecting airport(s)

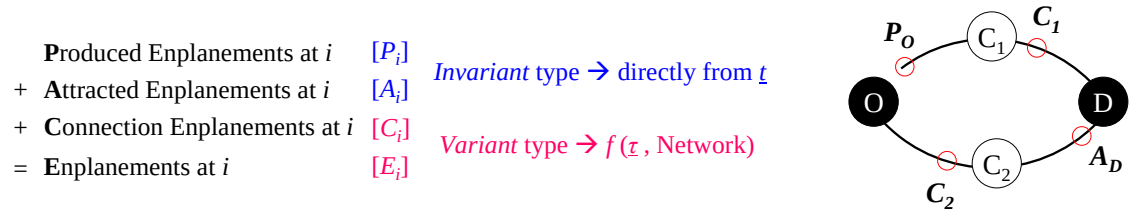


Figure 24: Notations for demand characterization. [Source: Lewe (2008)]

Clearly, C_i has a functional relationship with the passenger routings as well as with t_{ij}. In contrast, P_i and A_i, which are collectively categorized as generated enplanements (G_i), have an invariant nature with respect to the network. Figure 25 exemplifies the difference between t, τ, and E where there is a true O-D demand of size one from 'a' to 'b' on an airline network. In this case $t_{ab} = 1$ and $t_{ba} = 0$. If we assume that the traveler comes back home the next day and this daily demand occurs every day, on any day after the first day there is one person going from 'a' to 'b', and another person coming back from 'b' to 'a'. Therefore, the O-D demand becomes $\tau_{ab} = 1$ and $\tau_{ba} = 1$. If this O-D demand is serviced by nonstop flight (left figure), the network enplanements increases by two, i.e. $\Delta E = (E_{ab} = 1) + (E_{ba} = 1) = 2$. However if the O-D demand is serviced with a connection at a hub 'h' (right figure), the network enplanements increases by four, i.e. $\Delta E = (E_{ah} = 1) + (E_{hb} = 1) + (E_{bh} = 1) + (E_{ha} = 1) = 4$. This 'PACE' breakdown is the backbone of the proposed intrinsic demand model.

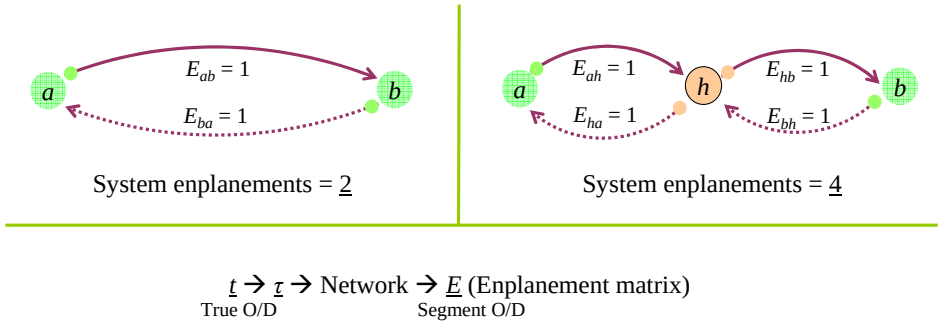


Figure 25: t, τ, and E on airline network. [Source: Lewe (2008)]

3.2.2 Available database

3.2.2.1 Origin and destination survey (DB1B)

Since the 1938 Civil Aeronautics Act, the federal government has collected and analyzed aviation data from commercial airlines. After the Civil Aeronautics Board Sunset Act of 1984, the Department of Transportation (DOT) became responsible for the handling of information on civil aeronautics. Currently, the aviation data that are submitted by U.S. air carriers are maintained by the Office of Airline Information (OAI) within Bureau of Transportation Statistics (BTS). Title 14, Code of Federal Regulations (CFR), PART 241 - UNIFORM SYSTEM OF ACCOUNTS AND REPORTS FOR LARGE CERTIFICATED AIR CARRIERS requires U.S. air carriers that provide scheduled passenger services with one or more aircrafts with the seating of 60 or more passengers to submit a quarterly origin and destination survey to the DOT. The source of this data is passenger tickets collected at the gate from the enplaned passengers. In this process, only tickets with the serial numbers ending in zero are selected. This process makes the total amount of data approximately 10 percent of the total enplaned passengers. Reporting air carriers compile this data in a prescribed format as illustrated in Figure 26.

The example itinerary in Figure 26 describes the information below.

1. Three passengers on this itinerary.

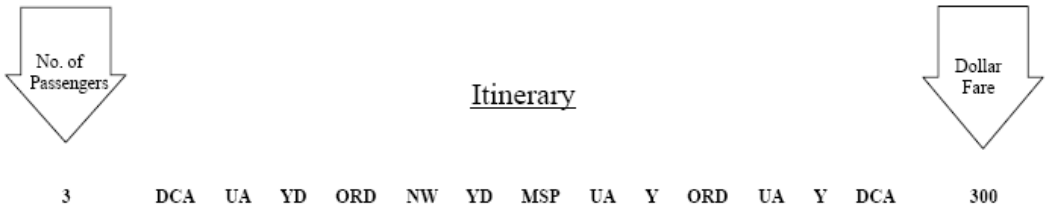


Figure 26: Illustration: A compressed line item in an O&D Survey. [Source: Bureau of Transportation Statistics (1998)]

2. Origin in Washington, D.C. National Airport (DCA) with United Airlines (UA) as the carrier and at a discounted coach fare (YD) to Chicago, Illinois O'Hare Airport (ORD).

3. Northwest Airlines (NW) as the carrier at a discounted coach fare to the destination of Minneapolis-St. Paul, Minnesota (MSP)

4. United Airlines as the carrier at a full coach fare (Y) when the passengers returned to Chicago.

5. United Airlines as the carrier at the full coach fare to Washington, D.C.

6. The total itinerary fare was $300 per passenger.

The Bureau of Transportation Statistics (BTS) provides the Airline Origin and Destination Survey (DB1B) database on their website (Bureau of Transportation Statistics, 2008). DB1B is a 10-percent sample of airline ticket information from reporting carriers and is composed of three databases, namely (1) DB1BCoupon, (2) DB1BMarket, and (3) DB1BTicket. The main difference between DB1BCoupon and DB1BMarket data is that the DB1BCoupon contains intermediate airport information for flights, while the DB1BMarket contains only the origin and final destination information obtained from their market-extracting algorithm. A snapshot of the download web site for DB1BCoupon is shown in Figure 27, and a sample of the data in DB1BCoupon is shown in Figure 28. The data includes origin, destination, and other details of itinerary of passengers with each record distinguished by itinerary ID

and market ID. Table 6 shows the enplanements change in the DB1B data. As seen in the table, the number of airports serviced has been reduced, suggesting that airline demand has been aggregated at more popular airports.

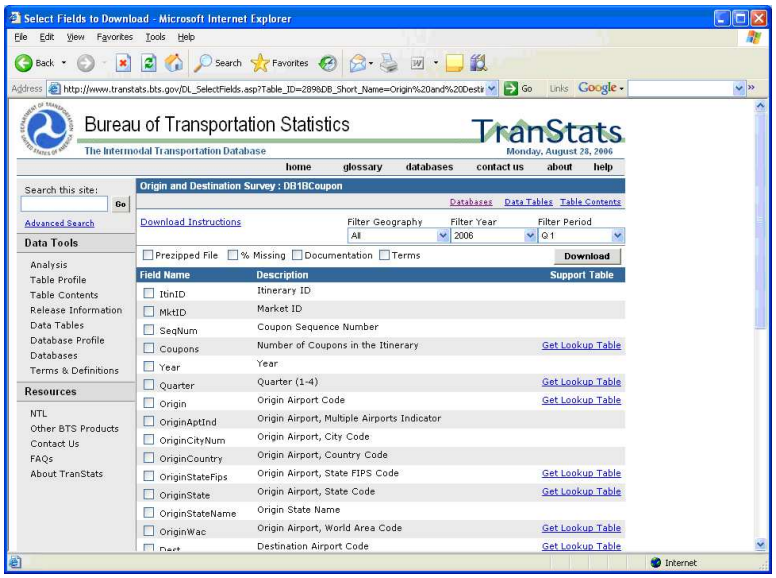


Figure 27: Snapshot - DB1BCoupon download site

If DB1B is accurate in their true O-D and transfer demand, passenger transfer rate can be obtained directly. In other words, DB1B data can be used to differentiate between visitors, connecting passengers, or returning residents. For example, if a record from DB1B shows no travel "break" for the arrival at an airport, then the traveler is a connecting passenger at the airport. If a record shows a "break" but it is not the concluding travel, then the traveler is a visitor to the airport. Finally, if a record concludes at the origin airport, then the traveler is assumed to be a local resident returning home (Detroit Metropolitan Wayne County Airport and University of Michigan, 2006). However, according to the audit reports by Bureau of Transportation Statistics (1998); Department of Transportation (2005), the O-D pair data from DB1B are neither accurate nor complete. Therefore, users should be careful in extracting O-D market information directly from it.

Unfortunately, DB1B does not have time information and the trip breaks are

ItinID	MktID	SeqNum	Coupons	Year	Quarter	Origin	OriginAptInd	OriginCityNum	OriginCountry	OriginStateFips	OriginState	OriginStateName	OriginWac	Dest	DestAptInd	DestCityNum
200613459216	200616255311	4	4	2006	1	CLT	0	17960	US	37	NC	North Carolina	36	SDF	0	51510
200613459217	200616255312	1	1	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	MHT	0	53971
200613459218	200616255313	1	4	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	ORD	2	18500
200613459218	200616255313	2	4	2006	1	ORD	2	18500	US	17	IL	Illinois	41	DSM	0	24180
200613459218	200616255314	3	4	2006	1	DSM	0	24180	US	19	IA	Iowa	61	ORD	2	18500
200613459218	200616255314	4	4	2006	1	ORD	2	18500	US	17	IL	Illinois	41	SDF	0	51510
200613459219	200616255315	1	2	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	ORD	2	18500
200613459219	200616255315	2	2	2006	1	ORD	2	18500	US	17	IL	Illinois	41	MKE	0	57850
200613459220	200616255316	1	4	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	ORD	2	18500
200613459220	200616255316	2	4	2006	1	ORD	2	18500	US	17	IL	Illinois	41	MKE	0	57850
200613459220	200616255317	3	4	2006	1	MKE	0	57850	US	55	WI	Wisconsin	45	ORD	2	18500
200613459220	200616255317	4	4	2006	1	ORD	2	18500	US	17	IL	Illinois	41	SDF	0	51510
200613459221	200616255318	1	1	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	PHL	1	69880
200613459222	200616255319	1	1	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	PHL	1	69880
200613459223	200616255320	1	1	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	PHL	1	69880
200613459224	200616255321	1	1	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	PHL	1	69880
200613459225	200616255322	1	1	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	PHL	1	69880
200613459226	200616255323	1	1	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	PHL	1	69880
200613459227	200616255324	1	1	2006	1	SDF	0	51510	US	21	KY	Kentucky	52	PHL	1	69880

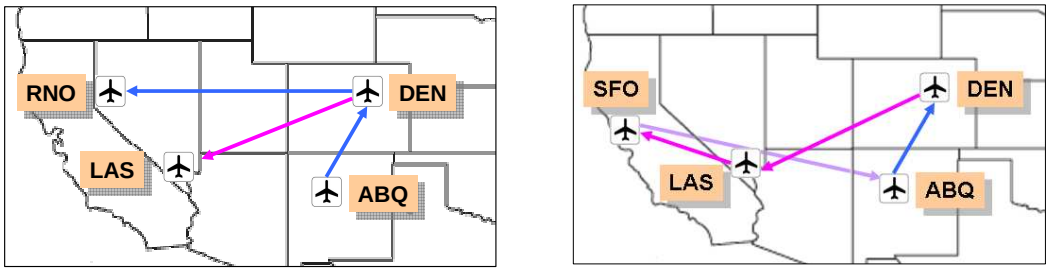
DestCountry	DestStateFips	DestState	DestStateName	DestWac	Break	CouponType	TkCarrier	OpCarrier	RPCarrier	Passengers	FareClass	Distance	DistanceGroup	Gateway	ItinGeoType	CouponGeoType
US	21	KY	Kentucky	52	X	A	US	16	ZW	1	X	336	1	0	2	2
US	33	NH	New Hampshire	14	X	A	US	ZW	ZW	1	X	820	2	0	2	2
US	17	IL	Illinois	41		D	UA	RP	ZW	1	X	286	1	0	2	2
US	19	IA	Iowa	61	X	A	UA	ZW	ZW	1	X	299	1	0	2	2
US	17	IL	Illinois	41		A	UA	UA	ZW	1	X	299	1	0	2	2
US	21	KY	Kentucky	52	X	D	UA	RP	ZW	1	X	286	1	0	2	2
US	17	IL	Illinois	41		D	UA	RP	ZW	1	Y	286	1	0	2	2
US	55	WI	Wisconsin	45	X	A	UA	ZW	ZW	1	Y	67	1	0	2	2
US	17	IL	Illinois	41		D	UA	RP	ZW	1	X	286	1	0	2	2
US	55	WI	Wisconsin	45	X	A	UA	ZW	ZW	1	X	67	1	0	2	2
US	17	IL	Illinois	41		A	UA	YV	ZW	1	X	67	1	0	2	2
US	21	KY	Kentucky	52	X	D	UA	OO	ZW	1	X	286	1	0	2	2
US	42	PA	Pennsylvania	23	X	A	US	ZW	ZW	1	X	576	2	0	2	2
US	42	PA	Pennsylvania	23	X	A	US	ZW	ZW	1	X	576	2	0	2	2
US	42	PA	Pennsylvania	23	X	A	US	ZW	ZW	1	X	576	2	0	2	2
US	42	PA	Pennsylvania	23	X	A	US	ZW	ZW	1	X	576	2	0	2	2
US	42	PA	Pennsylvania	23	X	A	US	ZW	ZW	1	X	576	2	0	2	2
US	42	PA	Pennsylvania	23	X	A	US	ZW	ZW	1	X	576	2	0	2	2
US	42	PA	Pennsylvania	23	X	A	US	ZW	ZW	1	X	576	2	0	2	2

Figure 28: Sample DB1BCoupon Data

Table 6: Enplanements from DB1B Data

	Ticket	Market	Coupon	#Airports in Ticket	#Airports in Market	#Airports in Coupon
1993	17,443,018	28,979,673	41,209,930	664	711	718
1994	19,860,005	32,634,801	46,196,314	613	687	701
1995	20,859,756	33,795,743	47,462,315	630	693	703
1996	22,451,307	36,367,042	50,741,430	631	683	693
1997	23,155,148	37,933,675	52,916,644	637	689	697
1998	22,534,464	38,823,126	54,269,640	640	690	697
1999	22,596,321	40,781,098	56,733,664	635	674	681
2000	23,778,182	42,863,761	59,350,552	616	649	659
2001	21,964,978	40,057,068	55,748,893	601	630	637
2002	21,154,107	38,735,926	54,291,718	570	606	609
2003	21,554,183	39,295,652	54,929,729	561	597	603
2004	23,557,758	42,617,212	58,725,814	518	554	556
2005	25,229,561	45,097,309	61,512,861	492	512	514

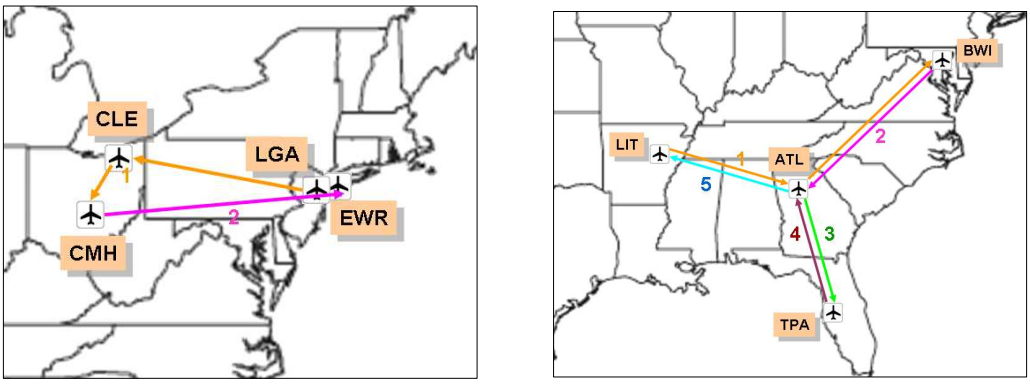
calculated according to their directional passenger concept which describes itinerary construction and circuity rules (Department of Transportation, 2005). Consequently, trips with multiple stops can be troublesome to determine their true origin and destination. One problematic example of this approach is shown in Figure 29. Figure 29a shows that the directional passenger concept used in DB1B always treats "Albuquerque (ABQ) → Denver (DEN) → Reno (RNO)" as a single directional passenger trip but divides "ABQ → DEN → Las Vegas (LAS)" into two single trips of "ABQ → DEN" and "DEN → LAS". In Figure 29b, the directional passenger concept divides a round trip "ABQ → DEN → LAS → San Francisco (SFO) → ABQ" into "ABQ → DEN", "DEN → LAS → SFO", and "SFO → ABQ".



(a) ABQ → DEN → LAS/RNO (b) ABQ → DEN → LAS → SFO → ABQ

Figure 29: Problems with DB1B circuity rules. [Example from Department of Transportation (2005)]

Missing time information in DB1B also generates unclear market information for multiple coupon trips. Some example travel data from DB1B is shown in Figure 30. Figure 30a shows a trip with two trip breaks, first of which at "CMH" and the second at "EWR". This trip is recorded as one-way trip since the origin ("LGA") and the destination ("EWR") are not the same. However, considering the proximity of the two origin and destination airports, this is essentially a roundtrip. Figure 30b shows a trip starting and ending at "LIT" with five trip breaks at "BWI", "ATL", "TPA", "ATL", and "LIT". Considering the fact that the trip breaks are computed using the directional passenger concept, this calculated O-D market can be debated.



(a) Practical roundtrip but recorded as oneway trip

(b) Trips can be interpreted in many ways

Figure 30: Unclear market information on DB1B. [Data source : DB1B]

As one can see from this example, direct usage of DB1B can result in misrepresentation of the true demand. There are also cases where round-trips are misrepresented as one-way trips if the travel involves ground transportation or separately purchased tickets. Since the true O-D demand cannot be obtained directly, a method is needed to estimate this measure.

3.2.2.2 1995 American Travel Survey

The 1995 American Travel Survey (ATS) was developed and conducted by BTS to collect information of long-distance travel made by people residing in the United States (Department of Transportation, 1999). The data includes over half a million person trips from 163 Metropolitan Statistical Areas (MSA) in the CONUS listing transport mode taken for the trips. This data can be used at the aggregate level but individual O-D level data has numerous errors that should not be neglected. Some of the errors were reported by Lim (2008).

3.2.2.3 Air carrier statistics (Form 41 Traffic) - T-100 data bank

All certified U.S. carriers and international carriers that have at least one point of service in the U.S. or its territories have to report monthly air carrier traffic information

on Form 41, Schedule T100. The data is collected by the Office of Airline Information (OAI) within the BTS and stored in the air carrier statistics database (Bureau of Transportation Statistics, 2009*a*). The coverage of the monthly data dated from 1990 to present.

This database is often used by various users to analyze traffic patterns and market shares as well as to generate reports on passenger, freight, and mails transported. T100 also contains number of passengers enplaned as well as aircraft types, and number of departures, making it possible to calculate carrier load factors. T100 is a widely accepted reliable benchmark for passengers in segmental trips. The database can be accessed through BTS web site (`http://www.transtats.bts.gov`). A snapshot of the download web site for T-100 Domestic (T100D) market data is shown in Figure 31.

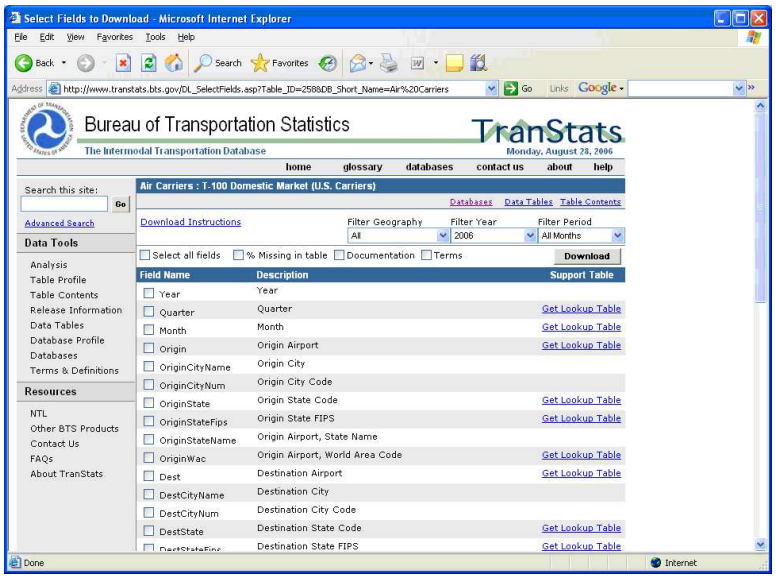


Figure 31: Snapshot - T-100 domestic market data download site

T100D is composed of two data, namely (1) T100D Market and (2) T100D Segment. T100D does not include passenger itinerary information and airfare, while DB1B does. Instead, the reporting carriers compile data from their market standpoint rather than the traveler's standpoint. Therefore, T100D Market data does not

correspond with passenger O-D demand. Figure 32 shows sample data collected from a T100D Market while Figure 33 shows sample data collected from a T100D Segment. Figure 34 illustrates the difference between segment and market data in T100D.

YEAR	QUARTER	MONTH	ORIGIN	ORIGIN_CITY_NAME	ORIGIN_CITY_NUM	ORIGIN_STATE_ABR	ORIGIN_STATE_FIPS	ORIGIN_STATE_NM	ORIGIN_WAC
2005	1	1	GSP	Greenville/Spartanburg, SC	33530	SC	45	South Carolina	37
2005	1	1	GSP	Greenville/Spartanburg, SC	33530	SC	45	South Carolina	37
2005	1	1	GSP	Greenville/Spartanburg, SC	33530	SC	45	South Carolina	37
2005	1	1	GUC	Gunnison, CO	34380	CO	8	Colorado	82
2005	1	1	HDN	Steamboat Springs/Hayden, CO	85730	CO	8	Colorado	82
2005	1	1	HDN	Steamboat Springs/Hayden, CO	85730	CO	8	Colorado	82
2005	1	1	HDN	Steamboat Springs/Hayden, CO	85730	CO	8	Colorado	82
2005	1	1	HNL	Honolulu, HI	36930	HI	15	Hawaii	2
2005	1	1	HNL	Honolulu, HI	36930	HI	15	Hawaii	2
2005	1	1	HNL	Honolulu, HI	36930	HI	15	Hawaii	2

DEST	DEST_CITY_NAME	DEST_CITY_NUM	DEST_STATE_ABR	DEST_STATE_FIPS	DEST_STATE_NM	DEST_WAC	AIRLINE_ID	UNIQUE_CARRIER	UNIQUE_CARRIER_NAME
AUS	Austin, TX	6100	TX	48	Texas	74	19790	DL	Delta Air Lines Inc.
DTW	Detroit, MI	24260	MI	26	Michigan	43	19790	DL	Delta Air Lines Inc.
ORD	Chicago, IL	18500	IL	17	Illinois	41	19790	DL	Delta Air Lines Inc.
ATL	Atlanta, GA	5852	GA	13	Georgia	34	19790	DL	Delta Air Lines Inc.
ATL	Atlanta, GA	5852	GA	13	Georgia	34	19790	DL	Delta Air Lines Inc.
CVG	Cincinnati, OH	19271	OH	39	Ohio	52	19790	DL	Delta Air Lines Inc.
SLC	Salt Lake City, UT	79071	UT	49	Utah	87	19790	DL	Delta Air Lines Inc.
ATL	Atlanta, GA	5852	GA	13	Georgia	34	19790	DL	Delta Air Lines Inc.
CVG	Cincinnati, OH	19271	OH	39	Ohio	52	19790	DL	Delta Air Lines Inc.
LAX	Los Angeles, CA	51410	CA	6	California	91	19790	DL	Delta Air Lines Inc.

UNIQUE_CARRIER_ENTITY	REGION	CARRIER	CARRIER_NAME	CARRIER_GROUP	CARRIER_GROUP_NEW	DISTANCE	DISTANCE_GROUP	CLASS	PASSENGERS	FREIGHT	MAIL
1260	D	DL	Delta Air Lines Inc.	3	3	957	2	F	20	0	0
1260	D	DL	Delta Air Lines Inc.	3	3	508	2	F	32	0	0
1260	D	DL	Delta Air Lines Inc.	3	3	577	2	F	21	20	0
1260	D	DL	Delta Air Lines Inc.	3	3	1301	3	F	453	0	0
1260	D	DL	Delta Air Lines Inc.	3	3	1340	3	F	514	0	0
1260	D	DL	Delta Air Lines Inc.	3	3	1202	3	F	616	0	0
1260	D	DL	Delta Air Lines Inc.	3	3	250	1	F	2242	0	120
1260	D	DL	Delta Air Lines Inc.	3	3	4502	10	F	14721	1136998	88533
1260	D	DL	Delta Air Lines Inc.	3	3	4433	9	F	5902	201906	9022
1260	D	DL	Delta Air Lines Inc.	3	3	2556	6	F	6927	899191	30168

Figure 32: Sample T100D market data

3.2.2.4 Summary

Finding intrinsic aviation demand starts from identifying credible sources of information. DB1B database is the only credible publicly available database that tracks passenger itineraries. There is also the 1995 American Travel Survey, however it was concluded in this work that these two sets of data project very different views of the travel market. Figure 35 compares the pdf of 'L-strip sum' from DB1B (left) and ATS (right). L-strip sum is a concept introduced to compare 3-dimensional matrix on 2-dimensional space by using Eq. (1).

$$L(i) = \sum_{j<i} (E_{ij} + E_{ji}), \qquad i \geq 2 \tag{1}$$

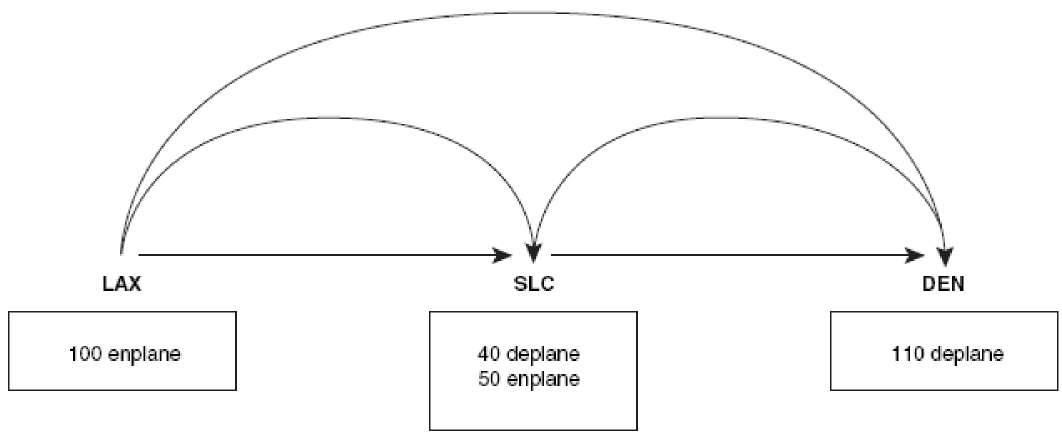
As such, they cannot be used together. Since real data are considered to be more reliable, DB1B was used for further investigation. Even so, during the course of this

YEAR	QUARTER	MONTH	ORIGIN	ORIGIN_CITY_NAME	ORIGIN_CITY_NUM	ORIGIN_STATE_ABR	ORIGIN_STATE_FIPS	ORIGIN_STATE_NM	ORIGIN_WAC	DEST	DEST_CITY_NAME	DEST_CITY_NUM	DEST_STATE_ABR	DEST_STATE_FIPS
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	BUF	Buffalo, NY	13742	NY	36
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	BWI	Baltimore, MD	7231	MD	24
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	BWI	Baltimore, MD	7231	MD	24
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	BWI	Baltimore, MD	7231	MD	24
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	BWI	Baltimore, MD	7231	MD	24
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	CLE	Cleveland, OH	19770	OH	39
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	CMH	Columbus, OH	20570	OH	39
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	CMH	Columbus, OH	20570	OH	39
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	DCA	Washington, DC	96460	DC	11
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	DCA	Washington, DC	96460	DC	11
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	DCA	Washington, DC	96460	DC	11
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	DCA	Washington, DC	96460	DC	11
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	DEN	Denver, CO	24091	CO	8
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	DEN	Denver, CO	24091	CO	8
2005	1	1	CVG	Cincinnati, OH	19271	OH	39	Ohio	52	DEN	Denver, CO	24091	CO	8

DEST_STATE_NM	DEST_WAC	AIRLINE_ID	UNIQUE_CARRIER	UNIQUE_CARRIER_NAME	UNIQUE_CARRIER_ENTITY	REGION	CARRIER	CARRIER_NAME	CARRIER_GROUP	CARRIER_GROUP_NEW	DISTANCE	DISTANCE_GROUP
New York	22	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	410	1
Maryland	35	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	430	1
Maryland	35	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	430	1
Maryland	35	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	430	1
Maryland	35	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	430	1
Ohio	44	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	221	1
Ohio	44	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	116	1
Ohio	44	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	116	1
District of Columbia	32	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	411	1
District of Columbia	32	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	411	1
District of Columbia	32	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	411	1
District of Columbia	32	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	411	1
Colorado	82	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	1069	3
Colorado	82	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	1069	3
Colorado	82	19790	DL	Delta Air Lines Inc.	1260	D	DL	Delta Air Lines Inc.	3	3	1069	3

CLASS	AIRCRAFT_GROUP	AIRCRAFT_TYPE	AIRCRAFT_CONFIG	DEPARTURES_SCHEDULED	DEPARTURES_PERFORMED	PAYLOAD	SEATS	PASSENGERS	FREIGHT	MAIL	RAMP_TO_RAMP	AIR_TIME
F	6	620	1	1	1	30765	100	100	0	12	87	59
F	6	614	1	1	1	41000	150	67	0	0	83	63
F	6	620	1	5	4	120340	400	232	1254	53	385	279
F	6	622	1	57	54	2763090	9864	4925	27877	5779	4955	3434
F	6	655	1	30	28	974400	3976	2204	7401	1395	2497	1797
F	6	620	1	92	85	2561595	8500	4346	6791	69675	5785	3620
F	6	614	1	1	1	41000	150	81	112	115	46	28
F	6	655	1	92	87	3027600	12354	5779	10900	18677	4634	2426
F	6	614	1	31	28	1148000	4200	2721	95	0	2490	1695
F	6	620	1	55	53	1593435	5300	4041	97	2838	4586	3260
F	6	655	1	62	59	2053200	8378	2656	330	672	4984	3569
F	6	656	1	1	1	39450	150	139	0	0	76	61
F	6	614	1	1	1	41000	150	126	1019	350	169	157
F	6	622	1	1	1	51705	183	57	1260	500	171	154
F	6	655	1	122	119	4141200	16898	11523	32016	93977	22684	19215

Figure 33: Sample T100D segment data



Key: DEN = Denver International Airport; LAX = Los Angeles International Airport; SLC = Salt Lake City International Airport.

Notes: **Nonstop segments** are represented by straight arrows, i.e., number of passengers transported between points (between takeoff and landings).
LAX to SLC: 100 passengers transported
SLC to DEN: 110 passengers transported
Onflight markets are represented by curved lines, i.e., where passengers are enplaned and deplaned on a flight.
LAX to SLC: 40 passengers
LAX to DEN: 60 passengers
SLC to DEN: 50 passengers
For a one-stop flight, the number of passengers would be the same as under segment and market.

Source: U.S. Department of Transportation, Office of the Secretary, *O&D Survey Reporting Regulations for Large Air Carriers: Code of Federal Regulations Part 241, Section 19-7* (Washington, DC: 1992).

Figure 34: Difference between segment and market data in T100D. [Source: Bhadra and Texter (2004)]

research, it was concluded that even the raw data from DB1B can be erroneous (Bureau of Transportation Statistics, 1998; Department of Transportation, 2005) and must be pre-processed before it can be meaningfully utilized to obtain true O-D and connecting enplanements data. Another public database that is useful for this study is the aforementioned T-100D which is a complete enumeration of the airline operations. Although T-100D does not track passengers, it served as a calibration datum at a later part of this work.

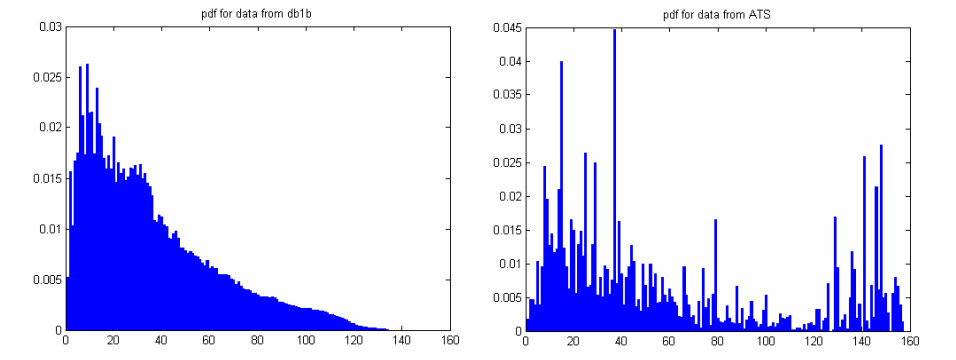


Figure 35: Difference between DB1B and 1995 ATS

3.3 Active design algorithm for airline network

This section presents an approach for active design algorithm of the airline network. The purpose of this research is to expand the NAS design space by incorporating network control variables into consideration. From an historical observations, it is hypothesized that the NAS is characterized as an evolutionary network with preferential attachment. Therefore, the proposed active design algorithm uses preferential attachment and introduces a concept of accelerated growth scheme. Additionally, the modeling effort will not be focused on building a detailed model but rather on capturing the general characteristic behavior of the NAS. This approach is necessary because the NAS is not an efficient system. Airline network has physical resources

that are costly to move around. Because of this resistance and highly uncertain environment which the network is exposed to, the NAS operates as a non-efficient system. While the purpose of modeling for an efficient system operating at optimal condition is to find a rigorous model for accurate market prediction, the purpose of modeling for an inefficient system is not so because of the behavior of not choosing optimal options. A modeler's role in this case should be to find a way to capture the general behavior. In this research, efforts have been given to build a concise model that captures general behavior of the evolutionary airline network. Other considerations for building a general airline operations model can be found in a paper by Rosenberger et al. (2002).

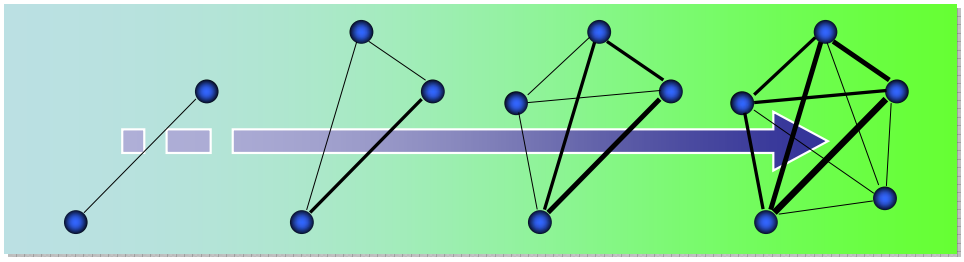
3.3.1 Accelerated evolution schemes

As observed in the previous section, the NAS shows many traits of preferentially attached network. Having established that the NAS is a non-efficient system, the goal of this research is to capture the behavior of the NAS rather than building an accurate model. Probabilistic choice model is introduced to serve this purpose and model the uncertainties involved in passenger routing. It reflects the situation when airlines and passengers do not choose the optimal options, which happens frequently in the NAS. In this respect, another sub-hypothesis is added to answer the research question 2-1 from Section 2.4 at this point as below:

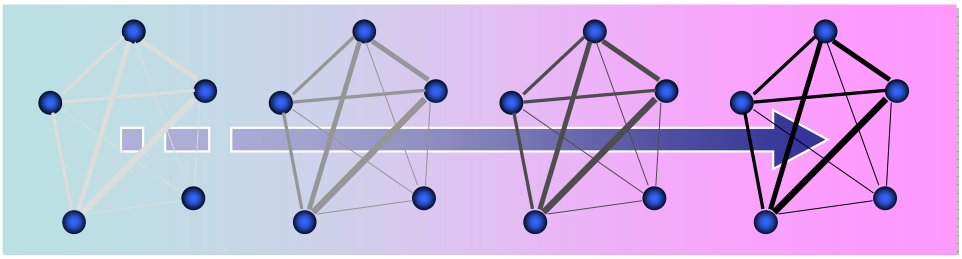
Hypothesis 2-1-3:
Probabilistic choice model can capture the non-optimal route choice behavior on the NAS.

In order to construct the target network from a baseline state, the question of how to evolve the network needs to be answered. For this purpose, two evolution schemes are introduced to simulate the evolution of the NAS. In the first scheme, the airports are added one by one to the network and connected to the pre-existing airports in the

network by preferential attachment. In this scheme, the airports introduced earlier gets more connections than the airports introduced later. The other evolution scheme starts with a certain number of airports and the network grows by a certain fraction of the total demand in each step. The first process is named as spatial expansion since it increases the number of nodes in the network, whereas the second scheme is named as chronological progression since it mimics the time-wise evolution of the network. These two basic schemes for accelerated evolution of NAS are shown in Figure 36.



(a) Spatial expansion: Adding a node (airport) one by one



(b) Chronological progression: Gradually increasing demand quantity

Figure 36: Two accelerated evolution schemes of the airline network model. [Source: Lewe (2008)]

It is hypothesized that these two evolution schemes are two representative characteristics of the NAS evolution. Therefore, another sub-hypothesis is added to answer the research question 2-1 from Section 2.4 at this point as below:

Hypothesis 2-1-4:
The growth of NAS can be emulated by spatial and chronological evolution scheme.

3.3.2 Route choice method

As O-D demand is served, routings for each respective demand has to be determined. The logit model detailed in Appendix. A.4 is used to compute the route choice. The process of using the logit model is described as follows:

1. The utilities associated with each routing are computed.

2. Using the logit model the probability of using a routing for transporting the O-D demand is represented as in the following equation.

$$P(i) = \frac{e^{U(i)}}{\sum_{z=1}^{n} e^{U(z)}}$$

where,

$P(i)$: Probability of selecting route i

$U(z)$: Utility associated with route z

n: Number of routings considered

The quantity assigned to each routing option is computed as a product of the O-D demand and the probability of each routing, $P(i)$. The route choice with the logit model can be performed with either utility or disutility associated with each route. In order to use the logit model with utility values, the profit and cost related to each option needs to be computed. There has been an effort by Brueckner et al. (1992) to link air fare and airline network structure, which can then be used to come up with utilities of airlines and passengers as a difference between benefit and cost for each side. And in order to use the logit model with disutility values, only the associated cost needs to be determined. The the utility or disutility values for an option i is computed as a weighted average of the airline and passenger cost, as in the following

equation:

$$Disutil_i = w \times AirlineDisutil_i + (1 - w) \times PaxDisutil_i$$

or

$$Util_i = w \times AirlineUtil_i + (1 - w) \times PaxUtil_i$$

where, w is a weight factor which designates whether the network formation is more favorable to the airline's benefit or to the passengers' benefit.

Another consideration is also given to the demand level for operation. In real airline network, if a given segment operation results in an average load factor below the breakeven point, a managerial decision needs to be made to avoid losses. The available options are 1) do not service the segment, 2) change airplanes to smaller size, or 3) service the demand with one or more stop operations. One or more stop operations are introduced where the existing segment demand alone is not sufficient but routing passengers from other segment can generate enough revenue to make the service economically viable. If the simulation is set to a threshold demand, some segments will not be served until the segmental demand reaches this threshold.

3.3.3 Segmental fleet selection

Planning for airline scheduling is a very large and complex problem, which is typically divided and optimized at each step. Airline flight scheduling and fleeting problem is an old problem that has been studied by many researchers (Yan and Tseng, 2002). Figure 37 shows the sequential process of airline schedule planning described in Lohatepanont and Barnhart (2004). Once the service plan is determined, only fleet routing needs to be planned. Airline routing plans are usually seasonal, covering mostly three months and typically with a weekly repetitive plan. Disturbances on the travel plan, such as canceled flights, can lead to the cancelation of other flights because of lack of availability of aircraft. In the route planning step, it is important to determine the minimum number of aircraft needed to cover the service plan.

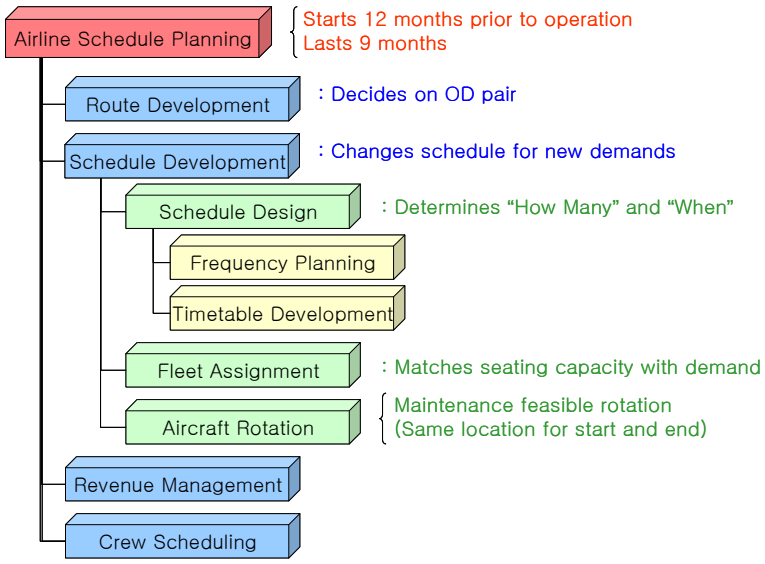


Figure 37: Typical airline schedule planning steps

In this research, one optimal aircraft type is assigned for each segment operation. Since the main purpose of this research is not in the development of a scheduling algorithm, a simplified method is used to obtain minimum aircraft and corresponding fleet for each segment. This segmental optimization with respect to aircraft type and quantity are different from general airline fleeting and scheduling addressed earlier. However, the difference is within the context of either globally or segmentally optimal solution. Therefore, it is assumed that the approach to find local optimal solution still captures the network formation trend.

The process of the segmental fleet selection is described in this section. If fixed cost and variable cost against characteristics of aircraft such as speed, range, and capacity are represented in equations, optimal vehicle and the frequency of the service in the segment can be obtained easily. One hypothetical case is shown in Figure 38. The plots of the fixed and variable cost can differ when viewed with different sets of independent variable values. Lederer and Nambimadom (1998) did a similar study where the variable cost per available seat-mile and the fixed cost per available seat-day are regressed against the aircraft size. In reality, aircraft have very diverse characteristics

which makes regression highly inefficient. Therefore, an alternative approach is taken with several representative aircraft represented by discrete sets of characteristics, A, B, C, D, and E in Figure 38. In the simulation, the best performing aircraft for each segmental operation is selected.

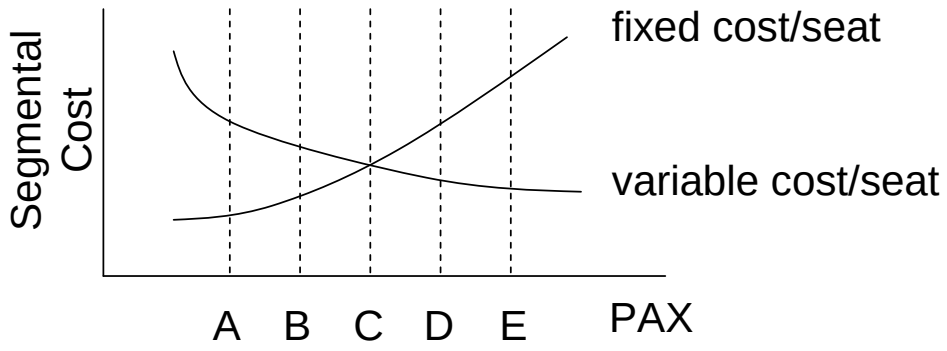


Figure 38: Rubberized vehicle in hypothetical case

There are also approaches to scale up the baseline model under expected growth factors to generate flight schedules (Viken et al., 2006*b*). However, growth factor based models are limited because they assume same resource usages, such as same load factor and aircrafts. Therefore, it is not a good option for the formulation of active design framework.

3.3.4 Simulation framework

In the study of transportation systems, closed-form analytical solutions normally do not exist because of a great deal of uncertainties and unknown relationships, which are hard to model. Therefore, an analysis of these system requires the credibility of modeling and simulation. In addition to combining several databases, a new true O-D based demand model and a evolution-based active design algorithm of the airline network model are required to address the hypotheses put forth in this research. Proper modeling environment is also a critical element in the design and management of this work.

Having considered the problems, research questions and hypotheses, a flexible environment with domain and analysis scalability is necessary. Domain scalability refers to the capability of incorporating different domain knowledges. Domain scalability is a valuable property when it is required to add a new system to the environment for which it was not originally designed. Analysis scalability refers to the flexibility of using different analysis programs and is a desirable characteristic of the environment when variable fidelity analyses are required in the different design phases. In an environment that contains rapidly evolving or revolutionary entities, an environment with domain scalability is demanded. In a design process that involves different phases of the product design, an environment with analysis scalability is demanded. In terms of software architecture an interface-based code is more advantages over a monolithic structured code. Therefore, design and implementation of the proposed models should be based on a highly scalable architecture. The framework will be built on modular architecture with interfaces so that the framework can be extended to incorporate new modules to be developed later or different level of fidelity modules can be composed and switched. The added advantage of having different fidelity-level codes is the capability to achieve reasonable simulation setup time and run time, depending on the importance of each module and availability of data.

Before Object-Oriented Programming (OOP) emerged, structured programming paradigm was popularly used to construct simulation environment. During this structural programming era, functional decomposition of the considered system was performed by taking a top-down design approach to divide a large system into sub operating units and then integrating them into a single monolithic synthesis code. This structured framework proved to be good for modeling a stationary system but due to the non-modular construction it had difficulties when there is a need for making changes. In many cases, small changes within a function can generate errors in many parts of the program because components are linked together in a monolithic code.

Emergence of OOP technique brought a new approach to systems modeling. Using OOP, the real system's properties and behaviors are contained in isolated models, i.e. classes. Having properties and related behaviors in isolated packages help designers to have a system-wide view since it resembles the real life system. The prominent benefits of object orientation in the development of a cooperative design environment lies in the encapsulation of behaviors. Encapsulating behavior becomes a very important part when there is a lot of of uncertainties on the actual construction of subcomponents. Hence, there is a lot to be gained by providing an architecture defined by interfaces. A cooperative environment that is defined by interfaces is also flexible enough to conveniently replace its components and incorporate future changes.

In summary, the three prominent benefits with using OOP can be described as below.

1. Encapsulation gives the benefit of modular construction and information hiding. Since each participant can build the environment on their own and hide important information, active participation of the stakeholders in the development of the design environment are encouraged. In this way the expertise of each discipline can be incorporated efficiently and the maintenance can be performed independently at the disciplinary level. Therefore, when changes to the model are needed, this can be done within the object limiting impact propagation.

2. Inheritance makes codes more reusable by letting subclasses share common behavior from superclasses. When there are many types of vehicles in the transportation network, some properties are common to all while others are unique to the each vehicle. By creating a parent object with common properties, any changes made to the parent object affect all child objects. Also, common interfaces can be enforced by using 'abstract class' and 'interface class'.

3. Polymorphism makes it possible to treat objects in the same family generically

and get different behavior from each object. Using polymorphism, it is possible to have different implementations of a certain functionality. As a result, calling the same function can perform different work for each entity.

CHAPTER IV

IMPLEMENTATION

The approach introduced in Chap.3 is implemented in this chapter. Specific setups of equations and parameter values are determined and the reason is explained.

4.1 Demand model implementation

4.1.1 Creation of sDB1B from DB1B database

As overviewed in Section 3.2.2, lack of time information make DB1B difficult to be used to deduce reliable true O-D demand. To avoid this problem, this research limits the scope to symmetric 2- and 4-coupon data with single-coupon trips paired-up according to the intrinsic true O-D demand ratios (i.e., t_{ij} to t_{ji}) obtained from the symmetric 2- and 4-coupon data. The final product after these pre-processing treatments is named the "sDB1B", which stands for Symmetric DB1B database. sDB1B for year 1997 accounts for 70 percent of the DB1B enplanements from the itineraries within the Continental United States (CONUS) and the total enplanements from DB1B are about 9.4 percent of the enplanements in T-100D segment data. sDB1B loses information on trips with more than two connections and multiple destination trips. However, the analysis showed that sDB1B sufficiently captures the characteristics of the original DB1B. By parsing all records in sDB1B, not only produced (P_i), attracted (A_i), and connection enplanements (C_i) at the airport level but also the enplanements at the segment level (E_{ij}) can be retrieved. A sample of the sDB1B is shown in Table 7. To make sure that the data maintains the characteristics of real world, E_{ij} from sDB1B was first compared to that from T-100D segment data. Figure 39 shows 3-dimensional depictions of enplanements matrix from each database

where the xy-plane represents a 470-by-470 matrix sorted by the size of airport enplanements in descending order.

Table 7: Sample records from sDB1B

ORIGIN	HUB	DEST	PAX
BDL	ATL	PFN	85
BDL	ATL	PHX	34
BDL	ATL	PIT	1
BDL	ATL	PNS	115
BDL	ATL	RDU	9
BDL	ATL	RIC	1
BDL	ATL	ROA	3
BDL	ATL	RSW	558
BDL	ATL	SAN	17
BDL	ATL	SAT	34
BDL	ATL	SAV	242

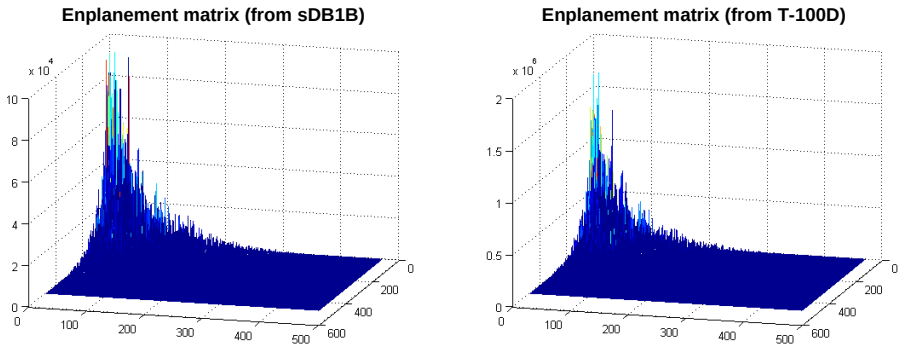


Figure 39: Enplanements comparison in 3D

As seen in Figure 39, it is difficult to gauge to what extent the two distributions are similar to each other. A numerical comparison was performed by converting the 3-dimensional matrix to a 2-dimensional form by using L-strip sum where the values in the enplanement matrix are summed up using Eq. (1), which was introduced earlier in Section 3.2.2. Figure 40 shows how the L-strip sum, $L(i)$, employed to compare the 3-dimensional data on a 2-dimensional space. The comparison of these two distributions shows that sDB1B represents the real data with good accuracy.

Here, Affinity Measure (AM) is defined in Eq. (2).

$$L(i) = \sum_{j<i}(E_{ij} + E_{ji}), \qquad i \geq 2 \tag{2}$$

$$AM = \frac{\sum_{i=1}^{N-1} |cdf_A(i) - cdf_B(i)|}{N-1}$$

where N is the number of airports in the network.

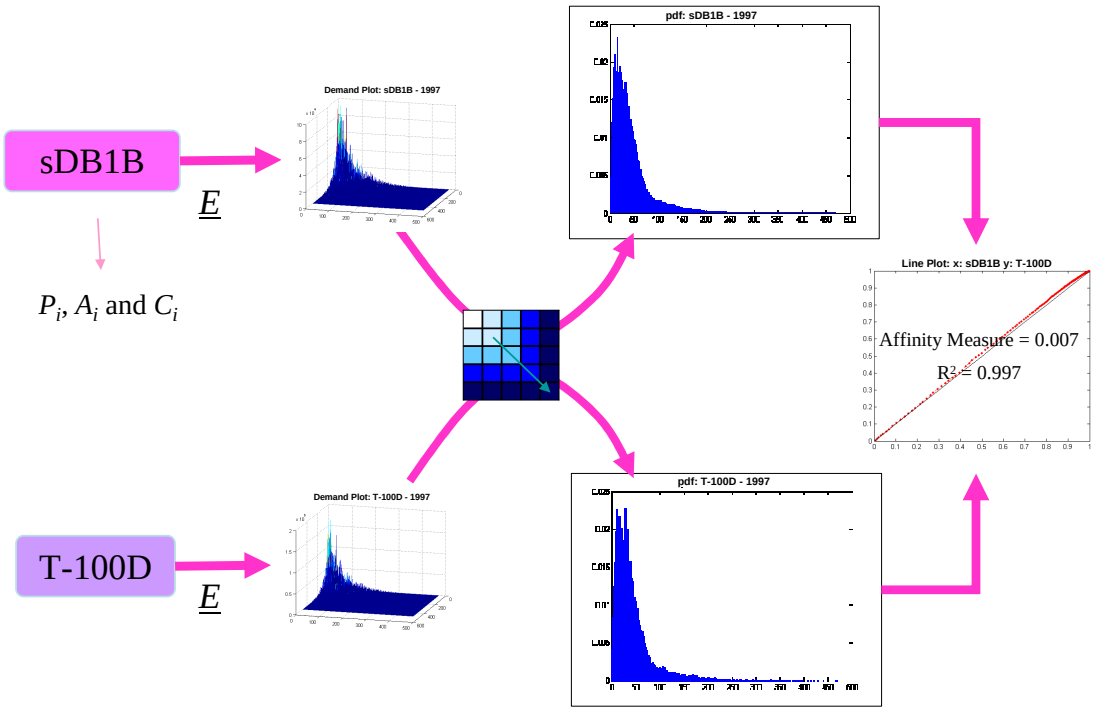
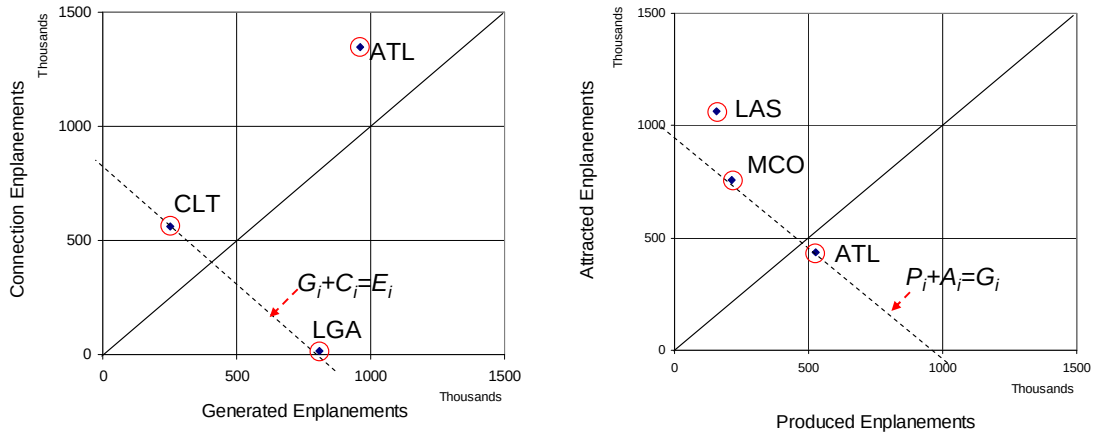


Figure 40: Enplanements comparison in 2D. [Source: Lewe (2008)]

4.1.2 Analysis of the demand data sets from sDB1B

Having created sDB1B, all of its data were processed with the PACE breakdown. The results gave a complete information on P, A, C for all airports. The two charts in Figure 41 demonstrate how this information can be analyzed for selected airports. In Figure 41a, airports above the 45-degree line have more connecting enplanements (C_i) than generated enplanements (G_i, recall that $G_i = P_i + A_i$), thus, playing the

role of hubs on the network. On the other hand, airports below the 45-degree line are the ones where majority of the enplanements represent true O-D demand. Also, the 135 degree line (dotted line) represents a line of constant total enplanements (E_i), indicating that airports close to an arbitrary 135 degree line have a similar number of total enplanements. Evidently, Atlanta (ATL) is depicted as a major hub with a large number of connecting enplanements that exceeds the number of generated enplanements. The figure also shows that while Charlotte (CLT) and LaGuardia (LGA) have similar total enplanements (close to the same dotted line), they take on very different roles in the NAS. Figure 41b shows the A_i vs. P_i graph. Airports above the 45 degree line have more attracted than produced enplanements, thus showing the characteristics of tourist cities. Airports on the same arbitrary 135-degree dotted line have similar number of generated enplanements. Most notably, the figure depicts Las Vegas (LAS) as a major tourist city that attracts a large number of air travelers. Also, even though Atlanta (ATL) and Orlando (MCO) have similar number of generated enplanements, MCO is evidently attracting more visitors than it is producing enplanements as compared to ATL which has a balanced number of attracted and produced enplanements.

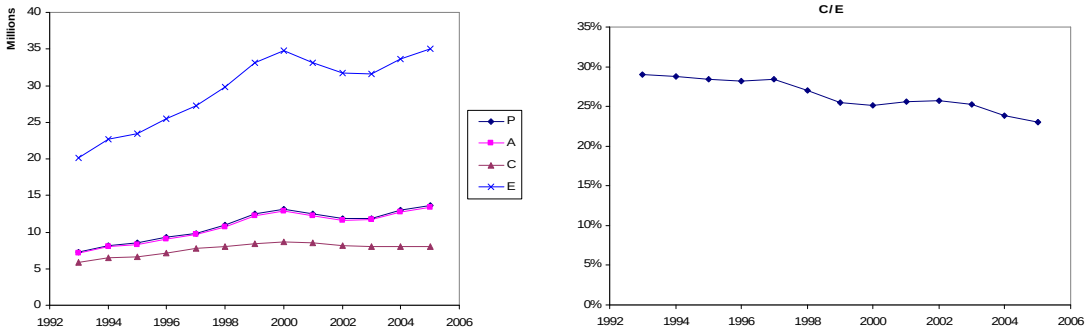


(a) C_i vs. G_i graph

(b) A_i vs. P_i graph

Figure 41: Demand characterization example plot with data from sDB1B, Year 2005

The PACE breakdown can also provide insights pertaining to the different forms of enplanements at the NAS-level. Figure 42a shows there has been a dramatic change in the composition of the different forms of enplanement after the year 2001, largely due to the propagating effects of September 11th. Figure 42b shows the interesting finding of a gradual decrease in the proportion of connection enplanements to total enplanements, suggesting that the NAS is moving towards a more P2P network. However, since sDB1B is a partial database of DB1B, which in itself is a collection of the total enplanements in the NAS, further investigations is necessary before any conclusive observations can be made.



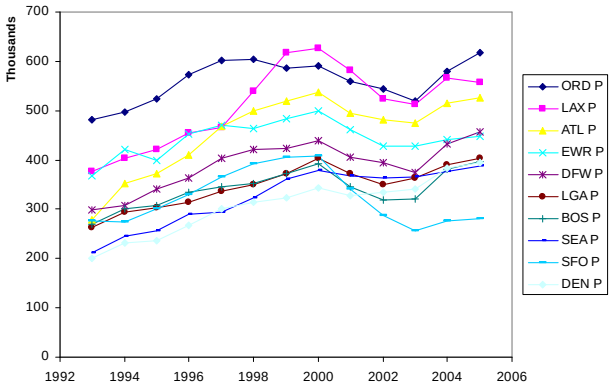
(a) PACE on NAS level

(b) Connection/enplanements on NAS level

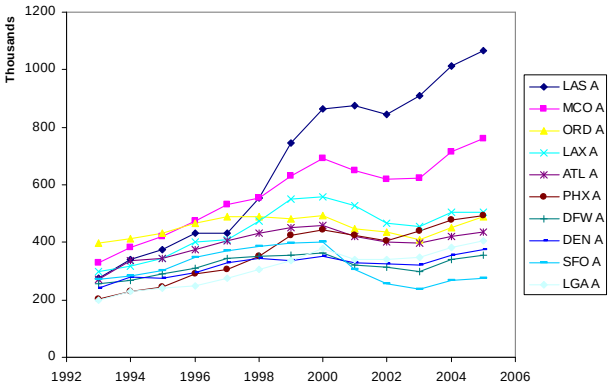
Figure 42: PACE analysis: NAS, from 1993-2005

After reviewing the NAS-level trends, airport-level characteristics were inspected as shown in Figure 43, with the produced, attracted, and connection enplanements for the top 10 individual airports in each measure. While produced and attracted enplanements are shown to gradually increase until year 2000 before experiencing a trend shift due to September 11th, connection enplanements does not reflect similar trends, which calls for further investigation of these relationships.

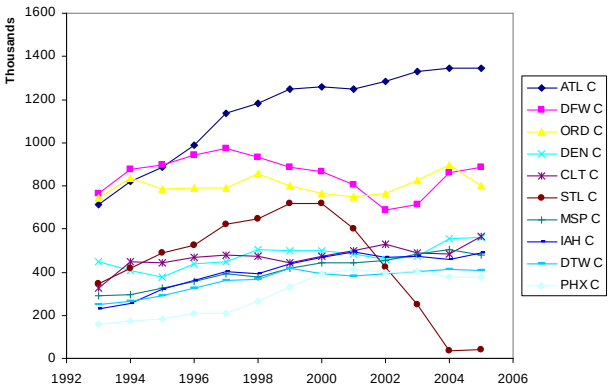
Figure 44 was obtained from sDB1B for the years 1993 through 2005 for the OEP35 airports which account for approximately 75 percent of the total enplanements in the U.S.. The trends of the PACE values in Figure 44 shows valuable information on the



(a) Trend of produced enplanements



(b) Trend of attracted enplanements



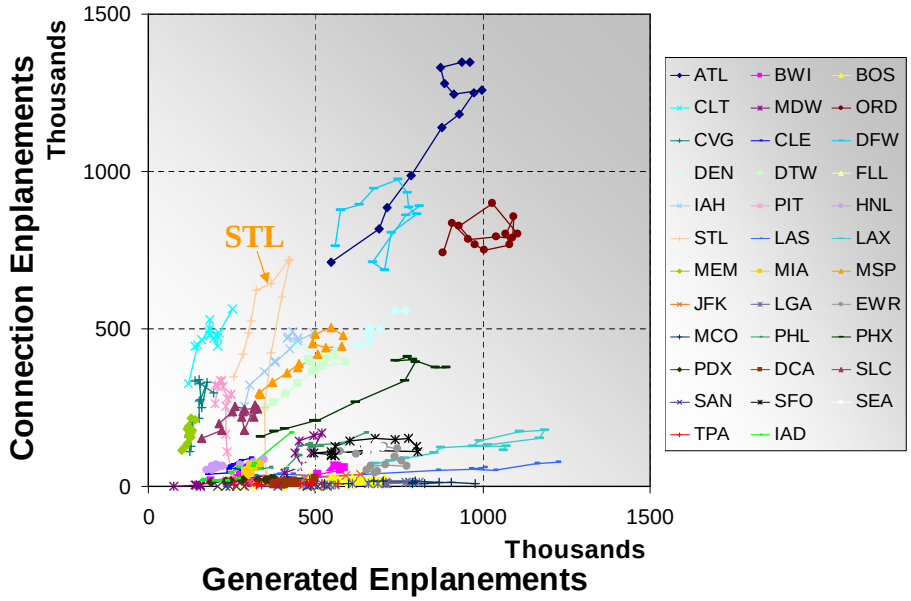
(c) Trend of connection enplanements

Figure 43: PACE analysis: top 10 airports, from 1993-2005

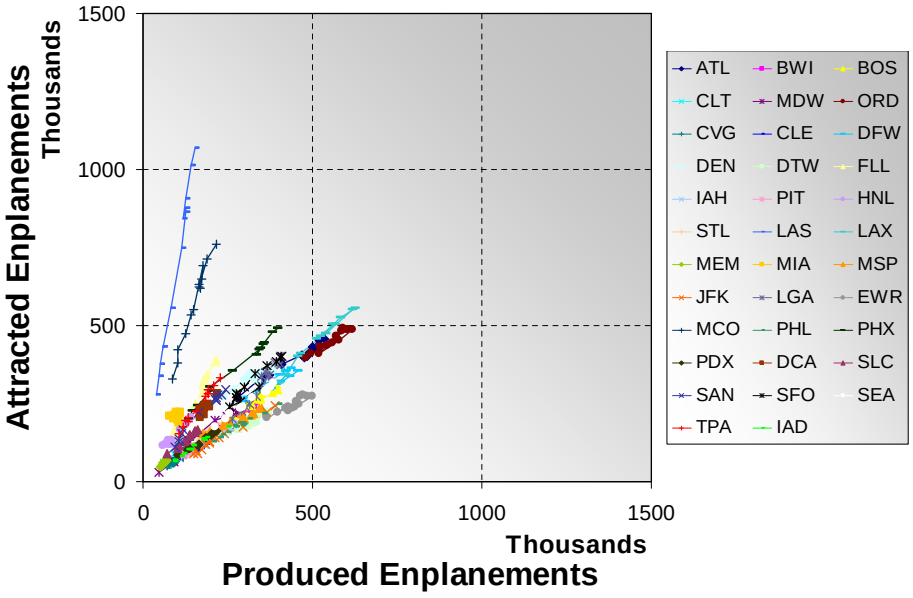
changing roles of the airports in the NAS. Figure 44a shows that there are no trends observed between C_i and G_i and the only meaningful observation is the dramatic drop in the "hubbing" activities at the St. Louis (STL) airport as a result of the dissolution of TransWorld Airlines (TWA). Since C_i is dependent upon network usage of the servicing airlines, which can vary on a short term basis, the lack of a general trend is not uncommon. Meanwhile, Figure 44b shows the relationship between A_i and P_i enplanements. In this plot, linear relationships are observed for all the airports, keeping the A_i/P_i at a nearly consistent ratio that increases only in total volume. The findings observed from the formulation of the sDB1B database suggest the possibility of intrinsic O-D demand model based on socioeconomic characteristics. Since the growth and declination of a region is usually gradual and attracted and produced trips (A_i/P_i) should be dependent on the socioeconomic characteristic of each region, it is hypothesized that these trends can be predicted if appropriate data are available.

4.1.3 Demand model hypothesis

Two observations can be made from Figure 44b. First, the trajectories of A_i and P_i for over a decade across all airports show very strong linear correlations. Practical usage of this finding is that once P_i is known then A_i can be solved for with assurance even if extrapolation of the linear trend is needed. In fact, 31 out of 35 OEP airports have the linear regression coefficient R^2 greater than 0.900 with the highest being 0.994. This linear relationship implies that modeling of A_i and P_i can be readily achieved as opposed to cases of having non-linear or even chaotic ones. Second, an intrinsic nature of an airport or its surrounding locales plays a role as observed by the different slopes of these linear correlations. From these observations, the following hypothesis is proposed: Model of A_i or P_i can be made by identifying socioeconomic characteristics of surrounding locales and finding their functional relationships with A_i or P_i. If proven true, the resulting aviation demand model is expected to improve



(a) Trends of C_i vs. G_i

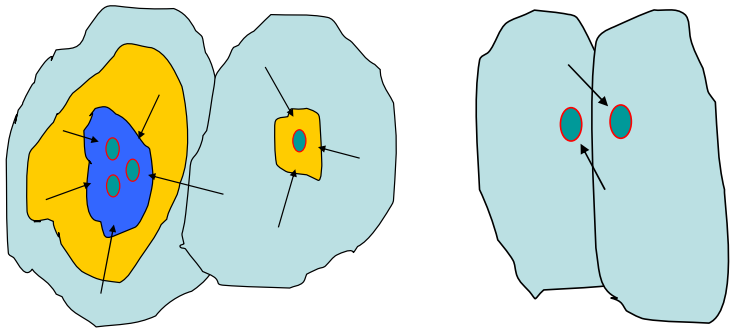


(b) Trends of A_i vs. P_i

Figure 44: Demand characterization of OEP35 using PACE, from 1993-2005

the predictive capability of today's demand estimation methods. To examine this hypothesis, a push and pull model was created. This push and pull model starts from the assumption that surrounding locales exert a certain influence to the target airport. The basic form of the equation should have attenuating factor that reduces the effect from each point mass as in a gravity field.

Also, the airports in close proximity should be combined as one since they serve the base demand of the region together. Therefore, categorizing by administrative boundaries should not be used. Figure 45 depicts this idea. Figure 45a shows that three airports on the left serves the more populated area while one airport on the right serves the less populated area. Figure 45b shows two airports which reside in different administrative region, such as MSAs, but closely located to each other, serving the same geometric region.



(a) Multiple airports serving regional demand

(b) Geometrically close airports, residing in different MSAs

Figure 45: Modeling level: Metroplex operation

Therefore, it is hypothesized that the influence from each county at a specific location has the form of Eq. (3) and (4).

$$X_i = \sum_k x_{ik} \tag{3}$$

$$x_{ik} = \beta_0 \frac{\prod_{j=1}^{n} x\_factor_{jk}^{\beta_j}}{\text{Dist}^{\beta_{n+1}}} \tag{4}$$

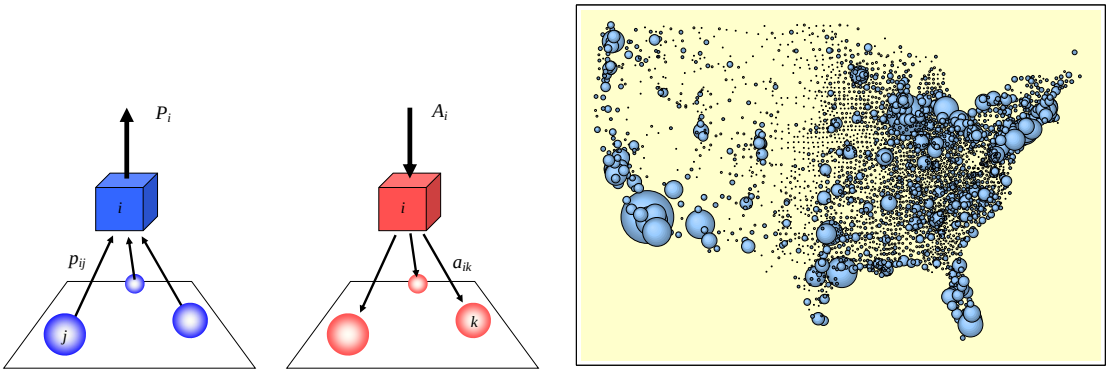
where X_i: Aggregate of individual influences from other locales (k) to locale i.

x_{ik}: Individual influence from other locales (k) to airport i.

β_0, β_j, β_{n+1}: Calibration coefficients.

x_factor_{jk}: Values of locale k for the selected factor j for the model.

In the above equations, the subscript k denotes a unit locale (county), x_factor represents an socioeconomic variable, and Dist is distance between the county and the location of interest, i. Figure 46a shows the pictorial description of the aforementioned model. Meanwhile, Figure 46b shows the distribution of population of all counties in the CONUS; each bubble representing a county located at its population centroid with the area being the relative population size.



(a) Influence from each element to P_i and A_i (b) 3,109 counties and population

Figure 46: Concept of the push and pull model. [Source: Lewe (2008)]

4.1.4 Aviation demand model on CONUS

For the implementation of the demand model, it was decided to include population in both A_i and P_i model. Additionally median household income was added as a x_factor for P_i model and subjectively weighted information on recreation industries, accommodation, and food services was added for A_i model, resulting in two x_factors for each. Hence, the Eq. (5) and (6) shows a generic form of the model for both A_i and P_i).

$$X_i = \sum_k x_{ik} \tag{5}$$

$$x_{ik} = \beta_0 \frac{(Pop_k)^{\beta_1} \times (x\_factor_k)^{\beta_2}}{e^{\beta_3 \times \delta_{ik}}} \tag{6}$$

where Pop_k: Population at locale k

β_0, β_1, β_2, β_3: Calibration coefficients.

δ_{ik}: Hypothetical distance between airport i and a unit locale k (county).

The model considers attenuating influence with exponential to a hypothetical distance, analogous to the gravity law. The δ_{ik} measures a distance between airport i and locale k by calculating $(h^2 + d_{ik}^2)^{1/2}$ with a calibration parameter h introduced to account for airport accessibility and a great circle distance d_{ik} in miles as a surrogate of route distance. Therefore the distance term (Dist) in Eq. (4) is replaced with $e^{\delta_{ik}}$ in this implementation.

The data for industry-specific economic data in county level can be obtained from Economic Census[1] collected by U.S. Census Bureau. Economic Census profiles the US economy every five years and reports are available for the years of 1992, 1997, and 2002. Due to the extensive amount of data, it is decided to base the model on year 1997, which is deemed a 'normal' year before the $September11^{th}$ attack. Population data at this level are only available for the year 1990[2] and 2000[3] and the linear interpolation was used to generate the data for the year 1997. Also, county population centroids are only available for the year 2000 and it is assumed that the population centroids did not change much. Due to the complexity of the problem,

[1]Available from `http://www.census.gov/epcd/www/econ97.html`, last accessed May 2009.

[2]Available from `http://www.census.gov/population/censusdata/90den_stco.txt`, last accessed May 2009.

[3]Available from `http://www.census.gov/geo/www/cenpop/county/ctyctrpg.html`, last accessed May 2009.

large- and medium-hubs in the CONUS in year 1997 which accounted for 87 percent of total enplanements in the US were selected for demand model development. These 66 airports are listed in Table 8. One thing to note is that the model combines metroplex airports together as one airport to account for the fact that those airports service the same set of base demand of the region. This is a necessary and preventive measure to make sure estimated production should not be arbitrarily increased by simply adding hypothetical airports. The considered metroplex airports are listed as below:

- New York, NY area: EWR, JFK, LGA
- Los Angles, CA area: BUR, LAX, ONT, SNA
- San Francisco, CA area: OAK, SFO, SJC
- Chicago area, IL: MDW, ORD
- Washington, D.C. area: BWI, DCA, IAD
- Dallas, TX area: DAL, DFW
- Houston, TX area: HOU, IAH
- Miami, FL area: FLL, MIA

Therefore, there are 53 nodes (8 metroplex + other 45 airports) in the demand model.

Table 8: Airports included in the research

ABQ	ATL	AUS	BDL	BNA	BOS	BUF	BUR	BWI	CLE	CLT
CMH	COS	CVG	DAL	DCA	DEN	DFW	DTW	ELP	EWR	FLL
HOU	IAD	IAH	IND	JAX	JFK	LAS	LAX	LGA	MCI	MCO
MDW	MEM	MIA	MKE	MSP	MSY	OAK	OKC	OMA	ONT	ORD
PBI	PDX	PHL	PHX	PIT	PVD	RDU	RNO	RSW	SAN	SAT
SDF	SEA	SFO	SJC	SLC	SMF	SNA	STL	TPA	TUL	TUS

4.1.5 Results of P_i and A_i

As a first step to building the production model, the x_factor in Eq. (6) for production should be determined and the median household income was the factor of choice. Median household income data for the year 1997 are available from the Small Area Income & Poverty Estimates (SAIPE) [4] recorded by the BTS. The following equation is then used to calculate individual influence on production from county k to airport i.

$$p_{ik} = \beta_0 \frac{(Pop_k/1000)^{\beta_1} \times (MedIncome_k/1000)^{\beta_2}}{e^{\beta_3 \times \delta_{ik}}} \tag{7}$$

This individual influence was summed up to generate P_i estimate, which was compared to P_i available from sDB1B. The coefficients β_0, β_1, β_2, and β_3 as well as h were varied until the result demonstrated a satisfactory fit. From this process, it was found that a majority of the airport models are well behaved but there are some airport models that do not have a good fit — admittedly, it was not anticipated that a simple model could describe the complex NAS. Later on, it was discovered that those problematic outliers could be categorized into two groups. The first group consists of airports that are sitting at a big metropolitan area where a much bigger metropolitan area is in its vicinity, typically 70 to 80 miles apart. The list of these airports, dubbed as aviation satellites, is shown in Table 9.

In the second group, two airports (ATL and DEN individually) and one group of metroplex airports (DFW and DAL together) are the only members, dubbed as aviation stars. The metropolitan area that has this type of airports, in contrast to the first group, has the opposite situation in general and is traditionally known as a major aviation hub. The locations of all airports are shown in Figure 47, with the squares, blue '+' symbols, and red colors representing the aforementioned metroplex airports,

[4] Available from `http://www.census.gov/hhes/www/saipe/`, last accessed May 2009.

Table 9: Aviation satellites and adjacent metros

Airport	Location	Adjacent Metro(s)
BDL	Hartford, CT	Boston, MA
		New York, NY
MKE	Milwaukee, WI	Chicago, IL
PBI	Palm Beach, FL	Miami, FL
PHL	Philadelphia, PA	New York, NY
		Washington, DC
PVD	Providence, RI	Boston, MA
		New York, NY
SAN	San Diego, CA	Los Angeles, CA
SDF	Louisville, KY	Cincinnati, OH
		Indianapolis, IN
SMF	Sacramento, CA	San Francisco, CA

aviation satellites, and aviation stars respectively. The remaining well-behaved airport models are depicted as black circles.

By eliminating these two outlier groups from the regression analysis, a very satisfactory fit was obtained when the coefficients were $\beta_0 = 0.56$, $\beta_1 = 1.00$, $\beta_2 = 1.50$, and $\beta_3 = 0.05$ and the parameter $h = 10$. The result is plotted in Figure 48 where the first group shows the estimated production being significantly higher than the actual values and the second group shows actual production being significantly higher than the estimated values.

The same process was repeated for the attraction model. After many attempts of trial and error, a linear combination of industry revenues collected from the economic census of 1997 grouped by the North American Industry Classification System (NAICS) code 713 (Amusement, gambling, and recreation industries), code 721 (Accommodation), and code 722 (Food services and drinking places) was selected for the x_factor for attraction ($attraction\_factor$) as in Eq. (8).

$$attraction\_factor = 3.5 \times NAICS\_713 + NAICS\_721 + NAICS\_722 \qquad (8)$$

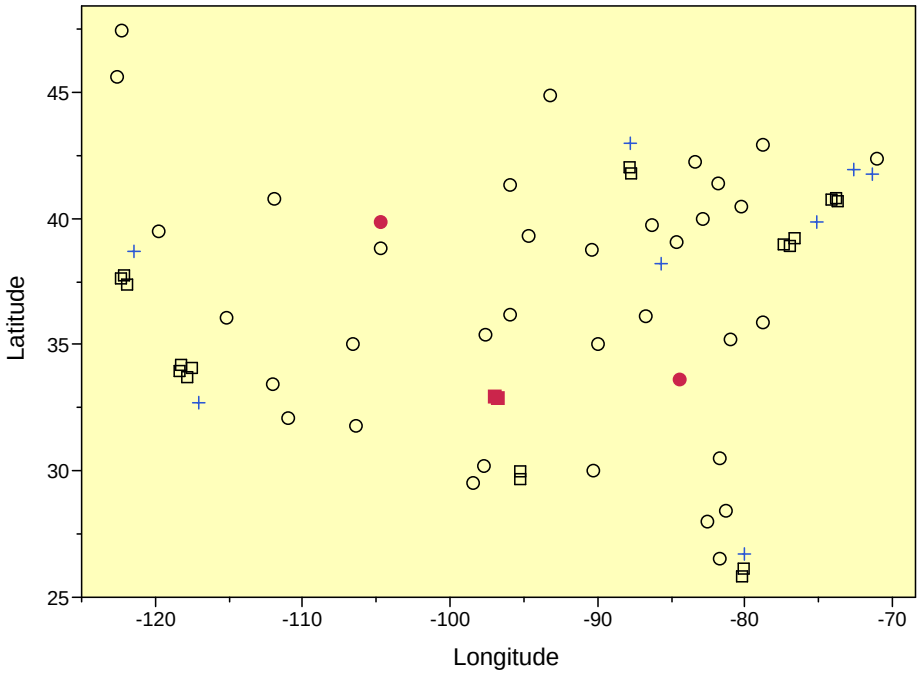


Figure 47: Locations of considered airports

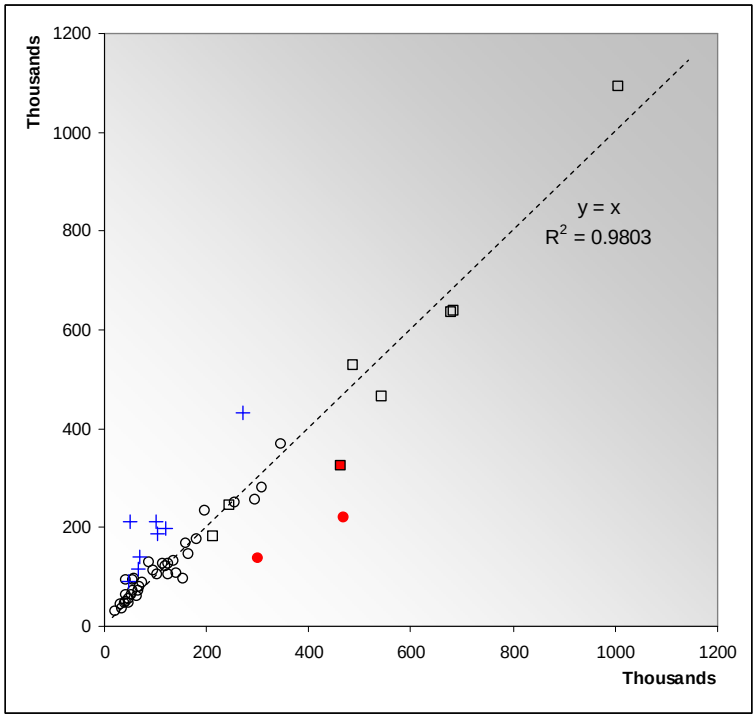


Figure 48: Production demand model result, x: P_i from sDB1B, y: P_i from production model

where 'NAICS_ddd': The annual sales figure in millions for the corresponding industry code.

As before, the two outlier groups were not included in the regression and the coefficients of the attraction model are $\beta_0 = 163.31$, $\beta_1 = 0.00$, and $\beta_2 = 0.90$. Note that β_3 and h were not altered to maintain consistency with the production model. Having $\beta_1 = 0.00$ suggests that the population was not a good variable to include since travelers may prefer less populated areas for their vacations. The result is shown in Figure 49, exhibiting a poorer result than the production model. This is partly due to the fact that much of the economic data for the attraction model were either unavailable or inaccurate: 53 percent of the county level data are missing for NAICS code 713, 31 percent for code 721, and 22 percent for code 713. It is also noted that the aviation stars and satellites are showing a similar trend on the graph as observed in the production model.

4.1.6 Two-factor FRATAR model

With P_i and A_i estimations from the push and pull model (in case one is mindful about a poor fitting of the attraction model, the linear correlation of A_i/P_i can be employed to obtain a more accurate A_i), a directional true O-D demand t_{ij} can be determined by a distribution algorithm such as the gravity or entropy model. More information on gravity and entropy model is included in Appendix A.3.

However, aviation demand data exhibits very high peak between specific popular city pairs — for example, between New York and Orlando, which makes it difficult to use them for demand distribution models. Therefore, growth factor models are well suited for air transportation demand distribution. The traditional growth factor models use one growth factor for each node because they deal with symmetric demand. On the other hand, the proposed demand model is based on directional true O-D

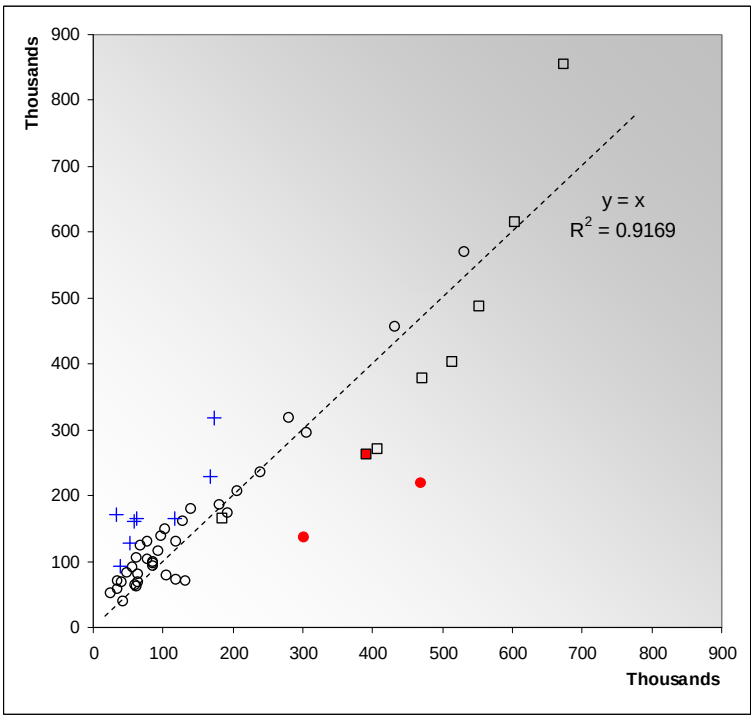


Figure 49: Attraction demand model result. x: A_i from sDB1B, y: A_i from attraction model

demand, with $t_{ij} \neq t_{ji}$. Therefore, two growth factors are needed for each node, i.e. factors for P_i and A_i.

The FRATAR model is the most popular growth factor model, which uses the ratio of trip generation quantity of the baseline and the future (Garber and Hoel, 2002). The FRATAR model relies on an existing symmetric O-D demand matrix which is equivalent to the τ matrix in this thesis. In each iteration, estimated trips from i to j becomes different than trips from j to i, i.e. $\tau'_{ij} \neq \tau'_{ji}$ and the pair-wise data are averaged to make them the same. Since it is a growth factor model, it cannot be used to predict demand between O-D pairs where no demand exists on the baseline. The FRATAR model has been widely used in air transportation research (Long, Lee, Gaier, Johnson and Kostiuk, 1999; Smith and Dollyhigh, 2004; Donohue and Zellweger, 2001) and hence, chosen for the necessary modification in this research.

4.1.6.1 Implementation of two-factor FRATAR model

Since t matrix is not symmetric and there has to be two growth factors for P_i and A_i, the traditional FRATAR model needs to be modified. The main reason for a two-factor FRATAR model is to have true O-D information in the prediction. When true O-D information is not needed, τ matrix can be used with traditional FRATAR process using predicted values of generated demands (G_is) as growth factors. The results will be the same. The focus is on the fact that the τ matrix can be generated from t matrix but not the other way round. Equation 9 shows the modified FRATAR with two factors for this research.

$$t'_{ij} = (P_i \times GP_i) \left[\frac{t_{ij} \times GA_j}{\sum_x t_{ix} \times GA_x} \right] \tag{9}$$

where

t'_{ij} = number of true O-D demand estimated from i to j

P_i = produced demand from i

GP_i = growth factor for P_i

t_{ix} = number of true O-D demand from i to x

GA_j = growth factor for A_j

In the equation, $(P_i \times GP_i)$ is an estimate for future trip generation of airport i. The denominator is an estimate for future attracted demand to all destinations from airport i. The numerator is an estimate for future attracted demand from i to j .

To illustrate the application of the growth factor model, the following example was devised. It uses the same t_{ij} matrix shown in Figure 7 in Chapter 1 reproduced here in Table 10. Forecasted demand increase for each airport for the future is shown in Table 11, where growth factors (GP_i and GA_i) are simply the ratio of future demand to current demand (Future-P_i/P_i and Future-A_i/A_i) and AP denotes the airport.

With the information shown in Table 10 and Table 11, the two-factor FRATAR

Table 10: Present true O-D demand t_{ij}

AP	A	B	C	D	P_i
A	-	60	50	10	120
B	40	-	30	30	100
C	50	60	-	20	130
D	90	70	80	-	240
A_i	180	190	160	60	-

Table 11: Predicted and current P_i and A_i

AP	P_i	Future P_i	GP_i	A_i	Future A_i	GA_i
A	120	144	1.2	180	198	1.1
B	100	110	1.1	190	247	1.3
C	130	182	1.4	160	224	1.4
D	240	312	1.3	60	72	1.2

proceeds as follows. First, the estimates for the next-step true O-D demand (t'_{ij}) is calculated using Equation 9. The calculations are shown below.

$$t'_{AB} = (120 \times 1.2)\left[\frac{60 \times 1.3}{60 \times 1.3 + 50 \times 1.4 + 10 \times 1.2}\right] = 70.20$$

$$t'_{BA} = (110 \times 1.1)\left[\frac{40 \times 1.1}{40 \times 1.1 + 30 \times 1.4 + 30 \times 1.2}\right] = 39.67$$

The same calculations are performed for all the elements in the matrix, resulting in the values shown in Table 12. As can be seen in the table, A_i does not match the future A_i, but P_i becomes the same as the future P_i at each iteration. To ensure convergence in subsequent iterations, GA_i is updated with the ratio of (future A_i) / A_i. The process iterates with updating GA_is until the convergence criteria is met. The final solution of this case is given in Table 13.

One thing to note is that the end result does not match 100 percent of the future A_i values. This is an inevitable problem of the growth factor model since one growth factor is used to scale multiple components of the matrix resulting in errors in the final results. Because the two-factor FRATAR only gives perfectly matching P_i results, the users of this approach should pay attention if the A_i results are within acceptable

Table 12: First estimate for true O-D demand t'_{ij}

AP	A	B	C	D	P_i	Future P_i	GP_i
A	-	70.20	63.00	10.80	144.0	144	1
B	39.67	-	37.87	32.46	110.0	110	1
C	63.76	90.42	-	27.82	182.0	182	1
D	102.28	94.01	115.71	-	312.0	312	1
A_i	205.7	254.6	216.6	71.1	-	-	-
Future A_i	198	247	224	72	-	-	-
GA_i	0.96	0.97	1.03	1.01	-	-	-

Table 13: Final result for true O-D demand t'_{ij}

AP	A	B	C	D	P_i	Future P_i	GP_i
A	0.00	67.69	65.40	10.91	144.0	144	1
B	37.94422	0.00	39.29	32.76	110.0	110	1
C	62.98649	90.01	0.00	29.01	182.0	182	1
D	98.93974	91.64	121.42	0.00	312.0	312	1
A_i	199.9	249.3	226.1	72.7	-	-	-
Future A_i	198	247	224	72	-	-	-
GA_i	0.99	0.99	0.99	0.99	-	-	-

range. Since the result from the demand modeling showed better performance with respect to P_i than A_i, the two-factor FRATAR was designed to generate exactly matching future P_i. Although the total errors can be improved if the accuracy of P_i can be sacrificed with different equation.

The two-factor FRATAR process scales up the current true O-D demand (t_{ij}). After the scale-up operation, the directional true O-D demand is summed up to generate $\tau_{ij} = t_{ij} + t_{ji}$, which are then fed into the network model to calculate enplanements on the network. The overall process is depicted in Figure 50.

4.1.6.2 Validation of two-factor FRATAR model for 2000 and 2005

The two-factor FRATAR model was tested to scale up the baseline t matrix. This validation followed the process depicted in Figure 51. The process is described as follows. True O-D matrix of year 1997 ($t1997$) was used as a baseline. P_i and A_i from year 2005 is inserted as scale factors to the two-factor FRATAR model with the

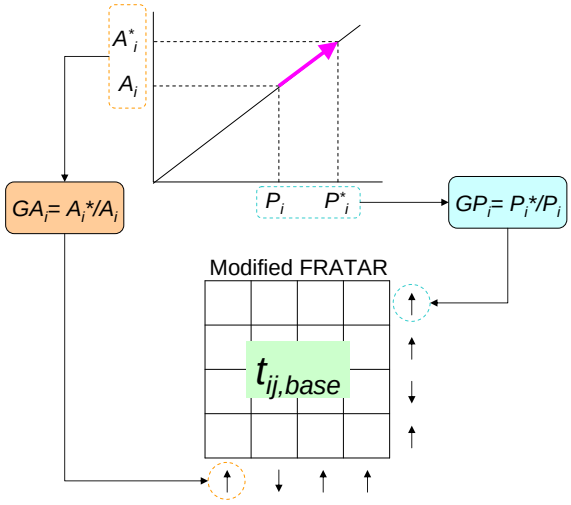


Figure 50: Two-factor FRATAR method with predicted P_i and A_i as growth factors

baseline t matrix. The resulting matrix is the predicted matrix for year 2005 ($t2005'$). Finally, this $t2005'$ matrix was compared with t matrix of year 2005 ($t2005$).

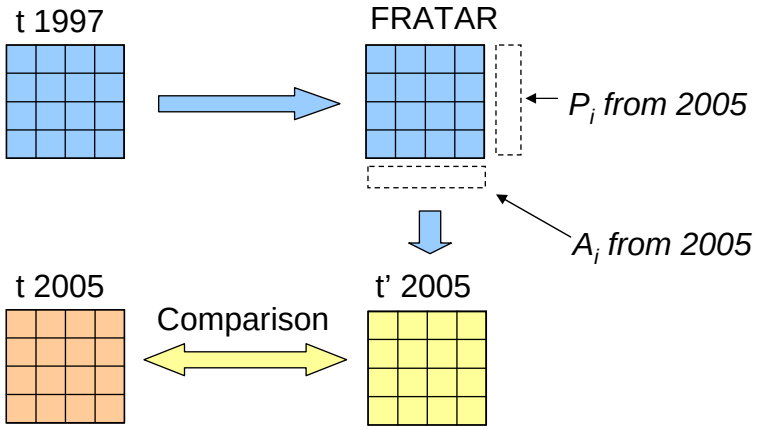


Figure 51: Two-factor FRATAR model validation process

In this study, scaling of the baseline t matrix to generate two target sets of year 2000 and 2005 was tested. The baseline t matrix is shown in Figure 52. The scaled ($t2000'$) and target ($t2000$) true O-D matrix (t matrix) is visualized in Figure 53. The result of the two-factor FRATAR applied to scaling year 1997 baseline to match year 2000 target t matrix shows almost perfect convergence of A_is from scaled matrix to the A_is from the target with the maximum individual difference of 2E-10. During the

iteration, the scaled matrix was set to have double precision while the target matrix is in integer. The sum of the squared error ($\sum_i (A_{i,FRATAR} - A_{i,2000})^2$) in this specific case was 1.14E-19. The subfigure 53c shows the Quantile - Quantile (QQ) plot of the L-strip sum of the scaled ($t2000'$) and target ($t2000$) matrix. The results show almost perfect scaling of the baseline to match the target. Figure 54 shows another case of the scaled($t2005'$) and target($t2005$) true O-D matrix (t matrix) validation. As observed in the investigation of the NAS in Section 3.1, NAS went through severe structure changes after year 2001. However, the proposed demand model based on P_i and A_i was able to scale-up the baseline t matrix of year 1997 to match the target year of 2005 almost perfectly. This is because the intrinsic demand maintains its characteristics as observed in the linear pattern, increasing the accuracy of the demand prediction if used as growth factors.

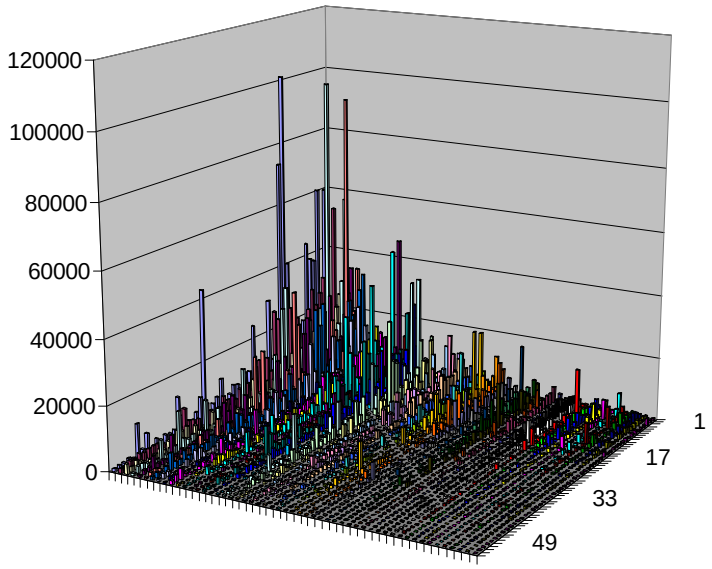
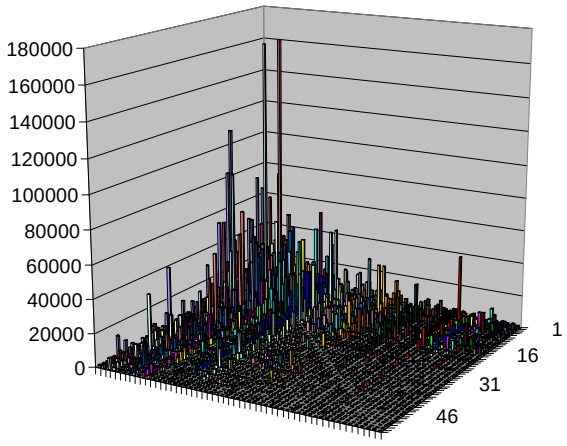
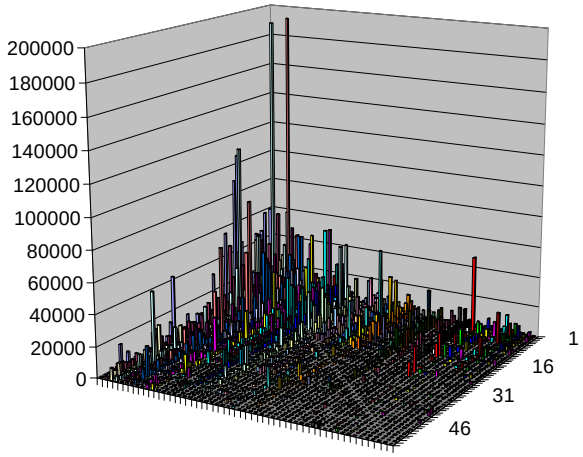


Figure 52: t matrix of year 1997

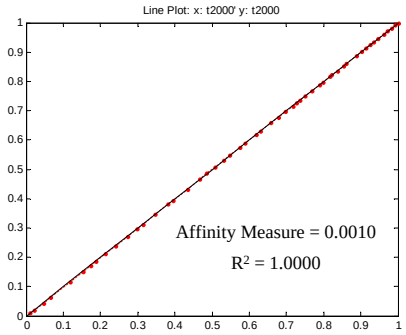
One thing to note is that the reason the two-factor FRATAR gives very good scaling here with respect to A_i is because P_i and A_i are highly correlated in this problem. When the two factors are uncorrelated, the resulting A_is from the scaled



(a) t matrix scaled from 1997 to 2000

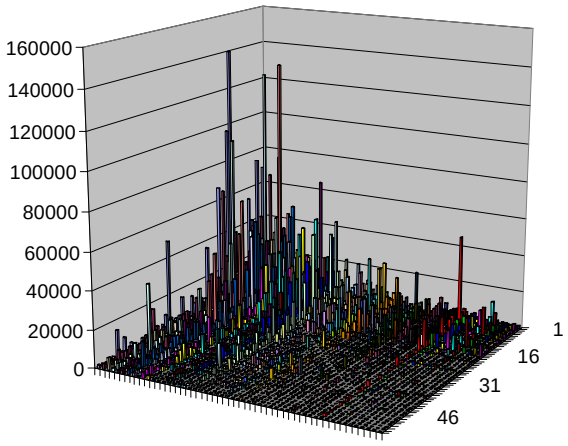


(b) t matrix of year 2000

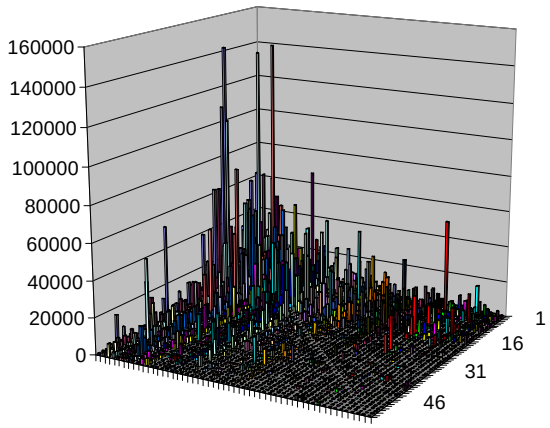


(c) Q-Q plot of scaled and target

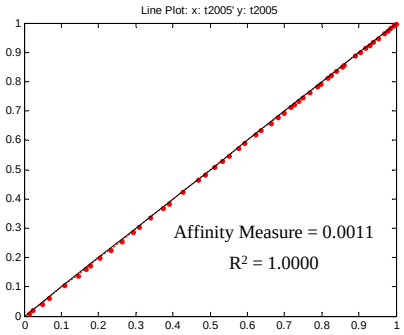
Figure 53: t matrix: scaled and target: 1997 → 2000



(a) t matrix scaled from 1997 to 2005



(b) t matrix of year 2005



(c) Q-Q plot of scaled and target

Figure 54: t matrix: scaled and target: $1997 \rightarrow 2005$

matrix can be quite different from the target values.

4.1.7 Summary

It was attempted to make an intrinsic O-D demand model that can be fed to the airline network simulation model to perform active design of it. Therefore, it was important to separate the variant and invariant demand. As a result, a new demand characterization technique named 'PACE' was tested and a demand model that can predict produced and attracted demand on the NAS was implemented. Even with the lack of socioeconomic data available for this model, the model still captures the prominent trends observed in the real world. However, the model was not able to generate reasonably good results for use as input to the active design algorithm. The main reason was the existence of the aviation stars and aviation satellites, which requires more in-depth analysis. However, further investigation could not be carried out at this point because of limited data.

Even though the demand model is not complete, this research brought a range of tangible benefits: a) the importance of intrinsic demand model in H&S network was emphasized, b) an airport's role on the air transportation network can be identified in more explicit way by the newly introduced PACE breakdown technique, c) modeling of P_i and A_i showed promising results, d) the process to be used for future demand prediction was illustrated with the two-factor FRATAR model. The scaled results demonstrated that the P_i and A_i can be used as growth factors with a baseline t matrix to formulate a demand prediction model that is robust against changes of the NAS structure.

4.2 Active Design Algorithm Implementation

The overall process is explained in the flow chart shown in Figure 55. The process proceeds as described below.

1. Initialization: Settings needed for the simulation is determined. The settings

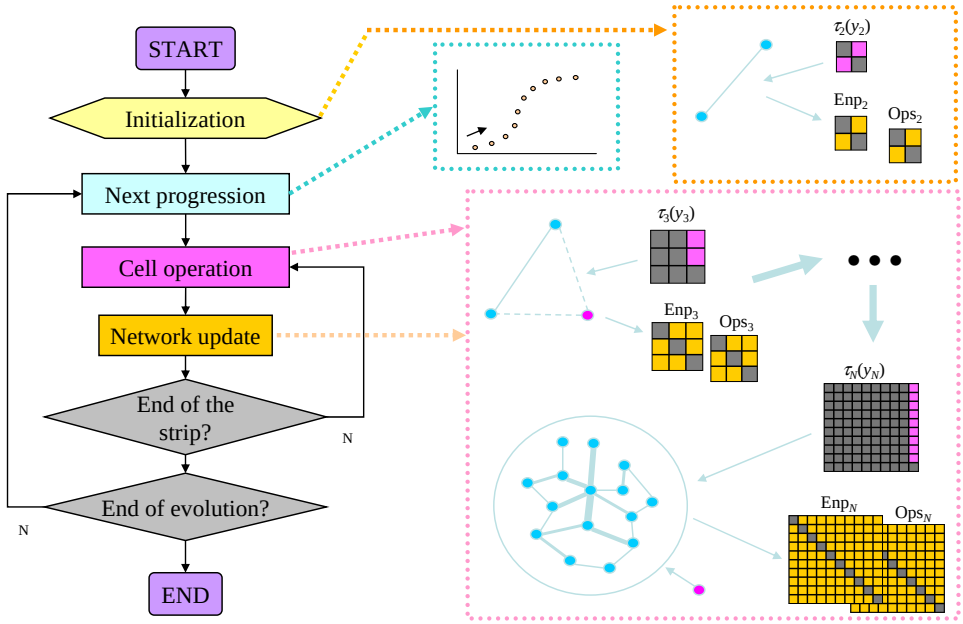


Figure 55: Active design algorithm flowchart

includes input file names, aircraft types to be used, values for hub-dials, and the evolution path.

2. Next progression: The program takes a next point on the evolution path — which determines the number of airports and the proportion of the demand that are needed to be put into the simulation — and prepare the inputs for the next cell operation.

3. Cell operation: A series of calculations called 'cell operation' are performed when a cell is selected. This is the core of the program where the trip demand of each cell (τ_{ij}) is distributed to appropriate routes.

4. Network update: Following the cell operation, the results of demand distribution into routes are updated in the program to reflect the calculation. Enplanements and operations matrix of the network is updated at this point.

5. End of the strip? : It is determined if the selected cell is the last cell to be considered for the evolution point. If it is the last cell then the process goes to

'End of evolution?'. Otherwise, the process goes to the next cell and the cell operation is performed for the cell.

6. End of evolution? : It is determined if the selected evolution point is the last one on the path. If it is the last evolution point then the program generates outputs and ends. Otherwise, the process goes to the next evolution point and the next progression is performed.

Detailed information on important processes are described in this section.

4.2.1 Accelerated growth

To test the sub-hypothesis SH2-1 and SH2-2, simulations will be performed in the 'evolution' space shown in Figure 56. The figure shows evolution trajectory which is a combination of both the spatial expansion and the chronological progression towards the final state. To simplify the evolution space, the demand fraction is discretized in the figure to have the same number of points as the number of airports but it does not have to be this way. Evolution path is one of the inputs that needs to be specified at the initialization phase. When a path that goes below or above the diagonal line as shown in the figure, their effects are very different. The former can be tried when a network is to be built with small number of the core airports first and mature the network to a certain point, and then add the rest into the matured network in a short time period. In many cases, networks observed in nature follows this hypothetical path. The latter path is the opposite of the former. It increases the network size very quickly and matures the already big network with longer temporal progression.

In each iteration, some amount of demand is added to the simulation. The amount is determined by the location of the point on the evolution path. When the next point is at the n/N on the chronological axis, the incremental demand is calculated by the following equation.

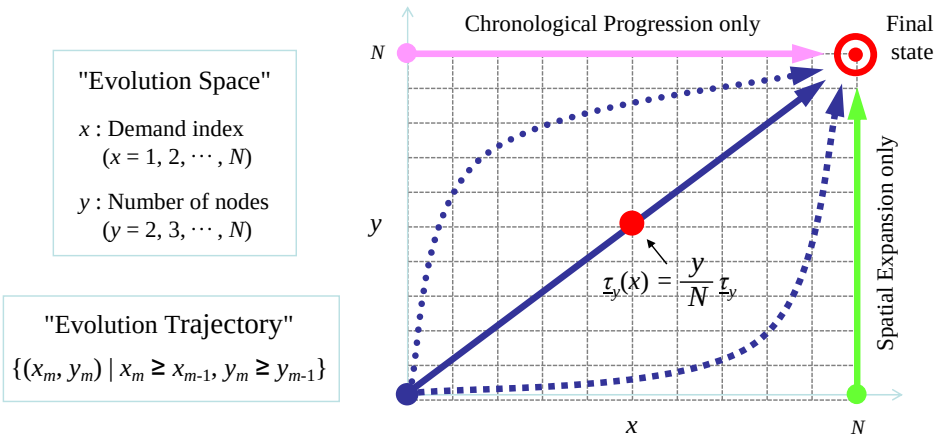


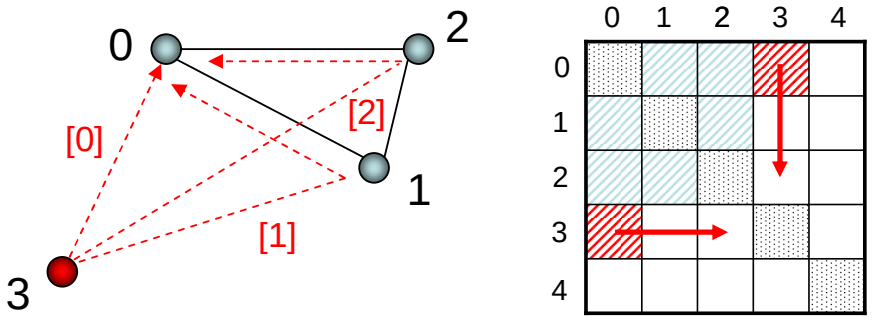
Figure 56: Evolution trajectory in evolution space. [Source: Lewe (2008)]

$$\Delta\tau_{ijn} = \tau_{ij} \times (n/N) - \sum_{k=1}^{k=n-1} \Delta\tau_{ijk}$$

where, n and k denote iteration point. This process is needed since incremental demand is added to the simulation as positive integer making decimal digits to be ignored. This process makes sure that n^{th} point has $\tau_{ij} \times (n/N)$ demand.

4.2.2 Cell operation

Cell operation is the basis of the active design algorithm. The same calculation is performed repetitively whenever O-D demand (τ_{ij}) or its partial set (τ'_{ij}) is distributed to routes. The cell operation is exemplified in Figure 57.



(a) Adding one node: Diagram (b) Adding one node: Matrix

Figure 57: Adding one node

The situation shown in this figure is where there are three airports in the network and the forth one is newly introduced. The cell operation starts with the cell (3,0), where the O-D demand between the introduced airport (AP_3) and the largest demanded airport in the network (AP_0), denoted as τ'_{30}, needs to be distributed to candidate routings. The following steps are taken to perform this job.

1. The disutilities (DU) of connection flights as well as nonstop flight for the selected O-D pair is calculated. In this example, there are three routing options as shown in Figure 57a. (In this research only nonstop and one-stop routes are considered as with the case of development of sDB1B. It was shown before that sDB1B captures the characteristics of the NAS.)

2. Probabilities of trip distribution is calculated based on the disutilities of the considered routes using the logit model. The probability of choosing routing i is represented as follows:

$$P(i) = \frac{e^{-DU(i)}}{\sum_{z=0}^{2} e^{-DU(z)}} \tag{10}$$

   where,

   $P(i)$: Probability of selecting route i

   $DU(z)$: Disutility associated with route z

3. O-D demand (τ'_{30}) is distributed on each route based on the probability $P(i)$. The following equation is used to perform this job.

$$\Delta\text{trip}_i = floor(\tau'_{30} \times P(i))$$

   where,

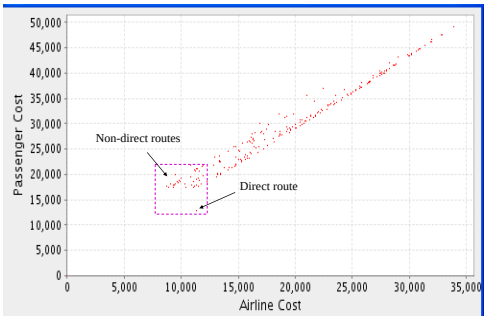
   Δtrip_i: Quantity of added demand on route i.

   $floor(z)$: A function that rounds the element z to the nearest integers less than or equal to z.

Because of the integer conversion with rounding down, there are cases where O-D demand (τ'_{ij}) is not completely distributed. This left-over demand is then added to the routes with high trip probability $P(i)$ one-by-one in a descending order until the remaining undistributed demand becomes zero. If there are multiple of routes with the same trip probability, higher demanded airports (which are closer to the top left corner on the trip matrix) are given priority.

When the τ'_{30} is distributed to the considered routes, the operation for the cell (3,0) is ended. The next cell (3,1) is selected and the same process continues. The iteration ends and the process goes out of the cell operation after the demand between the introduced airport (AP_3) and the least demand airport (AP_2) is distributed. It should be noted that when one cell is selected, for example (3,0), its symmetric cell (0,3) is also selected. This is because they describe symmetric demand that are needed to be transported in the opposite direction. Therefore, when a route is assigned a certain number of demand from airport i to airport j, there is an equal and opposite demand from airport j to airport i on the same route. Hence, the symmetric cells are marked together in Figure 57b.

Inclusion of very unlikely options in the probabilistic choice model can generate undesirable behavior since the model will generate probabilities for all included options. During this research, it was found that disutility values of routing options have only handful of cases that deserve consideration. It is well represented by the long and slender tail of alternative route options shown in Figure 58a. Therefore, it is desirable to select only reasonable routings for consideration. To do this, only Pareto as well as quasi-Pareto routes are considered when using probabilistic choice model. Figure 58b shows how quasi-Pareto options are selected. After Pareto options are identified, quasi-Pareto options are determined where the distance of them from any of the Pareto options are within the influence radius (r).



(a) Long tail of alternative routes

(b) Pareto and quasi-Pareto options

Figure 58: Selection of alternative routes for probabilistic choice model

4.2.3 Disutility modeling

It was determined that modeling cost as disutility is sufficient for modeling the growth of the network. The calculated disutility values are used to distribute the partial O-D demand (τ'_{ij}) to appropriate routes by employing the logit model. The basic characteristics of the probabilistic choice model used is described in Appendix A.4. The incremental combined disutility with respect to route i ($\Delta\text{Disutil}_i$) is represented with Eq. (11).

$$\Delta\text{Disutil}_i = [w \times \Delta\text{AirlineCost}_i + (1-w) \times \Delta\text{PaxCost}_i] \times \alpha \tag{11}$$

where $\Delta\text{AirlineCost}_i$: the incremental cost for serving the the portioned O-D demand (τ'_{ij}) through the route option i,

PaxCost_i: the cost associated with passengers for the route,

w: a parameter that puts weight on either one of them to reflect the importance of them on the formation of the network, w ranges from 0 to 1. A transport architect changes w to alter the network into more P2P by giving more weight to passenger disutility (decreasing w) or into more H&S by giving more weight to airline disutility (increasing w).

α: scale factor used to make the difference of disutility more or less discernable. (The higher, the more discernable.)

Airline cost (AirlineCost) is modeled differently from passenger cost. For a non-stop route, the cost of airline becomes the cost of operating aircrafts as follows.

$$\Delta\text{AirlineCost (O-D)} = \Delta\text{FixedCost} + \Delta\text{VariableCost (O-D)}$$

For a hub connection, the cost of airline becomes the following.

ΔAirlineCost (O-H_i-D) = [ΔAirlineCost (O-H_i) + ΔAirlineCost (H_i-D)] $\times(1 - AF_i)$

where AF_i is a hub attraction factor accounting for benefits the airline gets by using hub at i. The range of attraction factor is $[-\infty, 1]$, in which negative value means that there is penalty in making connection through the selected airport, positive means benefit, and zero means neutral. Figure 59 shows that hub attraction factor is applied to the incremental costs of origin-hub and hub-destination operations.

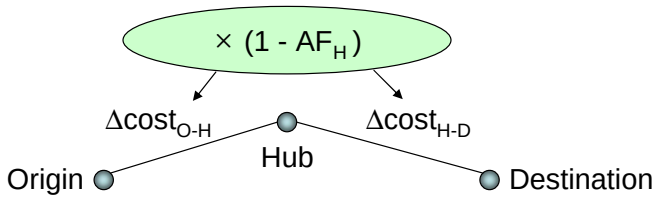


Figure 59: Description of hub attraction factor

Passenger cost (PaxCost) are calculated as product of travel time and passengers average monetary value of time (ValueOfTime). The travel time (TravelTime) is the total time spent for the itinerary including stopover. Equation for passenger cost is as follows:

ΔPaxCost (O-D) = $\#Pax \times \Delta$TravelTime (O-D) $\times$ ValueOfTime : for nonstop travel,

ΔPaxCost (O-H_i-D) = $\#Pax \times \Delta$TravelTime (O-H_i-D) $\times$ ValueOfTime : for hub connection,

where $\#Pax$ is the number of total passengers.

It should be noted that ΔAirlineCost is always positive or equal to zero, zero being that no additional aircraft operation is needed to transport the demand. However,

ΔPaxCost can be positive, zero, or negative. Negative ΔPaxCost happens when newly added demand justifies the usage of faster aircraft on a segment operation, for example from small turboprop to large jet, therefore reducing the total sum of the passenger travel time.

4.2.3.1 Normalization and de-normalization of disutility

If disutilities are used as raw values with the logit model, two problems may arise as described below.

- When the magnitude of the disutility is large, numerical instability can occur since e^{-DU} may become zero.
- When the difference of disutilities are small relative to their magnitude, the probabilities of choosing each route becomes indifferent. This becomes problem when the difference is not ignorable.

Therefore, it is desirable to have a scaling process to put data points in a certain range of values for the probabilistic choice model. A normalization process was performed for this purpose. There can be two distinct cases of extreme values as shown in Figure 60. The figure on the left shows two most dominant Pareto options each on one axis. There can be multiple Pareto options between these two points. The figure on the right is a case when there is a single Pareto option. The points on the figure is used as reference points for the normalization. The first step in the normalization is to eliminate unnecessary points to reduce computing time. Previously, the influence radius r was defined to include quasi-Pareto options for the choice model. Quasi-Pareto is defined as a point that is distanced from the Pareto options by no more than the influence radius (r). The quasi-Pareto options are identified on the normalized space. Therefore, the radius, r, is multiplied by the axis values of the reference points and the points outside of the boundary lines — any points that has larger values than $r \times x$ on x-axis or $r \times y$ on y-axis — are eliminated as shown in the figure.

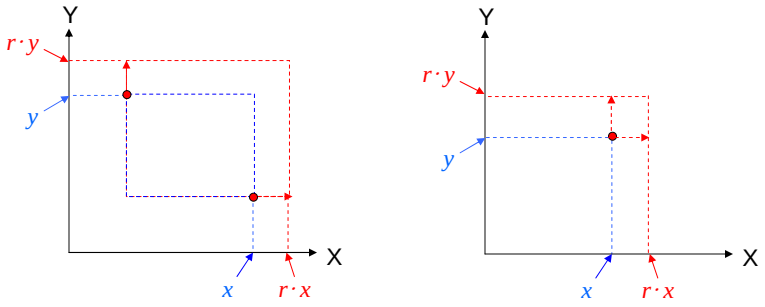


Figure 60: Cutoff of potential alternatives before Pareto and quasi-Pareto identification

After the elimination process, the normalization is performed as described in Figure 61. The figure on the left shows that when there are two reference points, they are relocated to (1,0) and (0,1) respectively. When there is a single reference point, it is located on (1,1) as shown on the figure on the right.

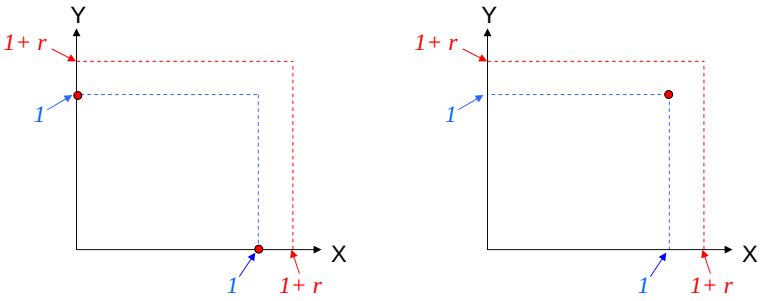


Figure 61: Normalization of disutility and consideration of quasi-Pareto options

The actual implementation of the case with two reference points are simpler than the case with one reference point. In this section x-axis represents airline disutility and y-axis represents passenger disutility. The equation for the normalization of the two reference case is below.

- $\text{disUtilX}_i = (\text{disUtilX}_i - \text{minX})/(\text{maxX} - \text{minX})$
- $\text{disUtilY}_i = (\text{disUtilY}_i - \text{minY})/(\text{maxY} - \text{minY})$

When there is a single reference point, airline disutility can only be positive or zero while passenger disutility can be positive, zero, or negative. However, passenger

disutility is always positive when airline disutility is zero. This is because having airline incremental cost of zero means that the same type and number of aircrafts are used to serve the added demand. Therefore, the incremental passenger time becomes always positive due to the added demand. This situation is depicted in Figure 62 with red dot being the extreme value point and blue dot being a neighboring one (left figure). The figure on the right shows possible areas of the location of the single Pareto point.

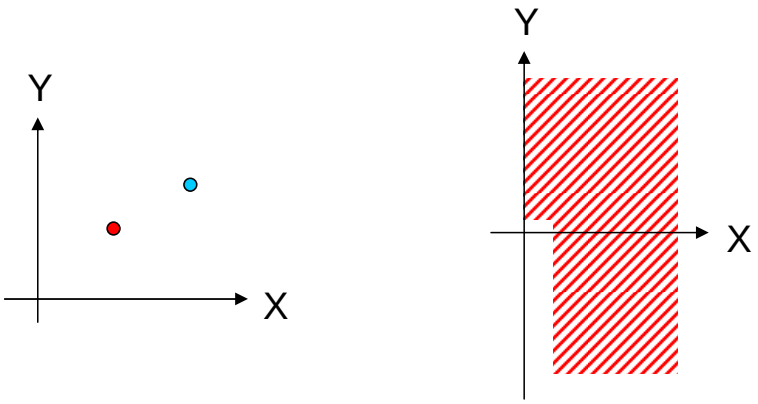


Figure 62: Possible location of single Pareto point

Note that airline cost can only be greater or equal to zero, where zero being the case the travel demand can be accommodated without adding more resources. Since the Pareto point will be scaled to have the x-axis value of 1, the following equation is used to proportionally scale other points.

1. if $\text{minX} > 0$ then $\text{disUtilX}_i = (\text{disUtilX}_i - \text{minX})/\text{minX} + 1$
2. if $\text{minX} = 0$ then $\text{disUtilX}_i = (\text{disUtilX}_i - \text{minY})/\text{minY} + 2$

When $minX = 0$, the first equation cannot be used. In this case $minX$ is replaced with $minY$ since $minY$ is always positive and it can also be used as a good reference value. The value of two is added to the equation to have disutility value of one when the routing i is the single Pareto point. In this case, the y-axis value is used as the same scaling base for both X and Y as described in Figure 63.

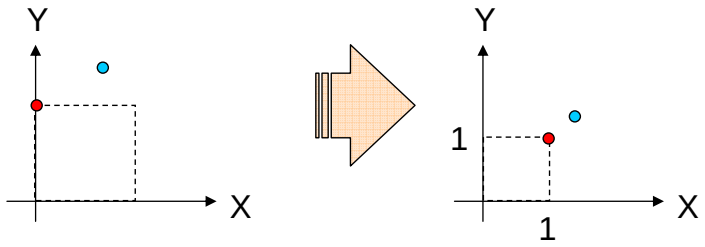


Figure 63: Same scaling for both axis when $minX = 0$

With respect to passenger disutilities, there are three cases to consider: 1) minY $>$ 0, 2) minY $= 0$, and 3) minY < 0. The following equations are used to proportionally scale other points.

1. if minY > 0 then $\text{disUtilY}_i = (\text{disUtilY}_i - \text{minY})/\text{minY} + 1$
2. if minY $= 0$ then $\text{disUtilY}_i = (\text{disUtilY}_i - \text{minX})/\text{minX} + 2$
3. if minY < 0 then $\text{disUtilY}_i = (\text{disUtilY}_i - \text{minY})/|\text{minY}| + 1$

The first and third cases can be represented together, i.e. if minY $\neq 0$ then $\text{disUtilY}_i = (\text{disUtilY}_i - \text{minY})/|\text{minY}| + 1$. When minY $= 0$, then minX is used as the scaling base for Y since it is positive in this case and it can also be used as a good scaling base. The reason why minX is positive when minY $= 0$ can be explained as follows. When minY $= 0$, which means that incremental passenger time with the added demand is zero, the aircraft type must have been changed to faster aircraft type on either one of the segments or both on the considered route. Therefore, the incremental airline cost should have positive value. So when minY $= 0$, the same scaling base is used for X and Y as in Figure 64.

After the normalization, the considered area gets mapped into a space with the value of one being at the reference point on each axis. Identification of quasi-Pareto options are performed in this normalized space for the ones within a circle with the influence radius r from the Pareto options as previously depicted in Figure 58b.

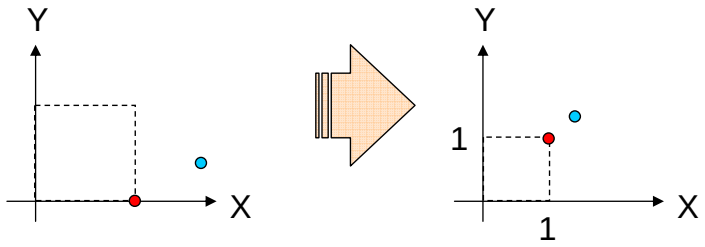


Figure 64: Same scaling for both axis when $minY = 0$

The de-normalization is performed to return the disutility value to relative dollar amounts. This process is necessary because using normalized disutility removes the information on the dollar difference between options. This relative dollar difference of the disutilities is used as the input for the logit model to calculate trip distribution. This dollar difference is different from the difference of absolute dollar value of the disutility. Example in Figure 65 depicts this process. The de-normalization into relative dollars is done by dividing each coordinate value with the maximum absolute values of the dollar-value on both axis. The figure on the left shows initial two Pareto values of (50, 500) and (250, 50) mapped to (0,1) and (1,0) respectively for the determination of quasi-Pareto options. The figure on the right shows the location of the two points after the de-normalization. In this example, all the four coordinate values were divided by 500. By dividing all the dollar values by the absolute maximum, it maintains relative difference between considered options. The de-normalization is performed for both Pareto and quasi-Pareto options.

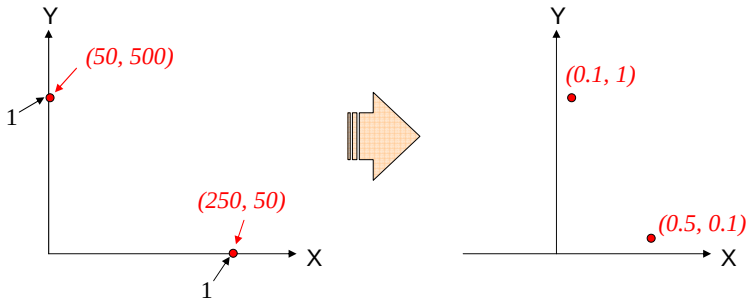


Figure 65: De-normalization of disutility into relative scale

The actual implementation considers the same cases as in the process of normalization. The equation for the de-normalization of the two reference case is below.

- $\text{disUtilX}_i = (\text{minX} + \text{disUtilX}_i \times (\text{maxX} - \text{minX}))/\text{maxAbs}$
- $\text{disUtilY}_i = (\text{minY} + \text{disUtilY}_i \times (\text{maxY} - \text{minY}))/\text{maxAbs}$

Where maxAbs is the maximum of absolute values of all X and Y coordinate elements.

When there exist single reference point, there are two cases for X and three for Y to consider as in normalization. The equation used in each case of X is the following.

1. if $\text{minX} > 0$ then $\text{disUtilX}_i = (\text{minX} + (\text{disUtilX}_i - 1) \times \text{minX})/\text{maxAbs}$
2. if $\text{minX} = 0$ then $\text{disUtilX}_i = (\text{minY} + (\text{disUtilX}_i - 2) \times \text{minY})/\text{maxAbs}$

The equation used in each case of Y is the following.

1. if $\text{minY} \neq 0$ then $\text{disUtilY}_i = (\text{minY} + (\text{disUtilY}_i - 1) \times |\text{minY}|)/\text{maxAbs}$
2. if $\text{minY} = 0$ then $\text{disUtilY}_i = (\text{minX} + (\text{disUtilY}_i - 2) \times \text{minX})/\text{maxAbs}$

4.2.4 Airline model

4.2.4.1 Representative aircrafts

The active design algorithm performs segmental optimization with respect to the selection of airplane types. It is decided that each segment considers five options of representative aircrafts. To select these five aircraft types, the information from T100D segment database in year 2005 was analyzed. In this process, the extracted aircraft types from T100D segment database were ordered by number of passengers and also by operations. The top 95 percent of aircraft types both in total number of operations or passengers were selected for the categorization, which are 23 aircraft types. Direct Operating Costs (DOC) for these 23 aircraft types were obtained from air carrier financial data of 'Schedule P-52' from US DOT From 41 in year 2005. The resulting list of the 23 aircrafts types are in Table 14. The table shows group

identification number which were categorized by range and capacity of the aircraft as shown in Figure 66. The speed on the table was obtained by by taking weighted average with number of operations from T100D segment data within each category.

Table 14: Considered aircrafts for selection of five aircraft classes

Group	AC_ Type	Description	Range	Pax	Speed	#Oper.	Avg. DOC/AirHours
1	461	Embraer Emb-120 Brasilia	887	30	345	206,730	1,108
1	456	Saab-Fairchild 340/B	920	30	300	288,017	1,204
1	483	Dehavilland Dhc8-100 Dash-8	1,174	37	310	228,343	1,373
2	629	Canadair Rj-200er /Rj-440	1,496	45	530	1,246,960	1,892
2	628	Canadair Rj-100/Rj-100er	1,496	45	530	237,261	2,030
2	676	Embraer-140	1,899	44	517	147,071	1,527
2	631	Canadair Rj-700	1,939	70	544	301,950	2,002
2	674	Embraer-135	1,956	37	517	147,699	1,577
2	675	Embraer-145	1,956	50	517	913,942	1,555
3	608	Boeing 717-200	1,646	110	570	204,609	3,280
3	640	Mcdonnell Douglas Dc-9-30	1,882	115	570	139,732	4,035
3	650	Mcdonnell Douglas Dc-9-50	1,882	135	558	54,345	3,917
3	620	Boeing 737-100/200	2,140	104	485	67,555	3,799
4	655	Mcdonnell Douglas Dc9 Super 80/Md81/2/3/7/8	2,359	155	504	731,147	3,769
4	617	Boeing 737-400	2,491	159	485	112,879	3,453
4	619	Boeing 737-300	2,612	128	485	847,870	2,931
4	622	Boeing 757-200	2,620	183	528	480,259	4,399
4	616	Boeing 737-500	2,764	123	485	195,402	3,216
4	614	Boeing 737-800	2,789	160	531	282,050	3,382
5	694	Airbus Industrie A320-100/200	3,452	150	517	452,811	3,104
5	612	Boeing 737-700/700lr	3,872	140	514	522,363	2,160
5	698	Airbus Industrie A319	4,258	124	517	381,145	2,865
5	626	Boeing 767-300/300er	4,546	269	530	64,383	5,436

To calculate airline disutility, the information of aircrafts on range, capacity, speed, fixed cost (FC), and variable cost (VC) are needed.

For assessing VC either a generally accepted practice or a publicly available airline economic database used. One of the most well-known practice for VC calculation is a method introduced by the Air Transport Association (ATA) for estimating DOC of airplanes, which was published in 1944 followed by some revisions in 1948, 1955,

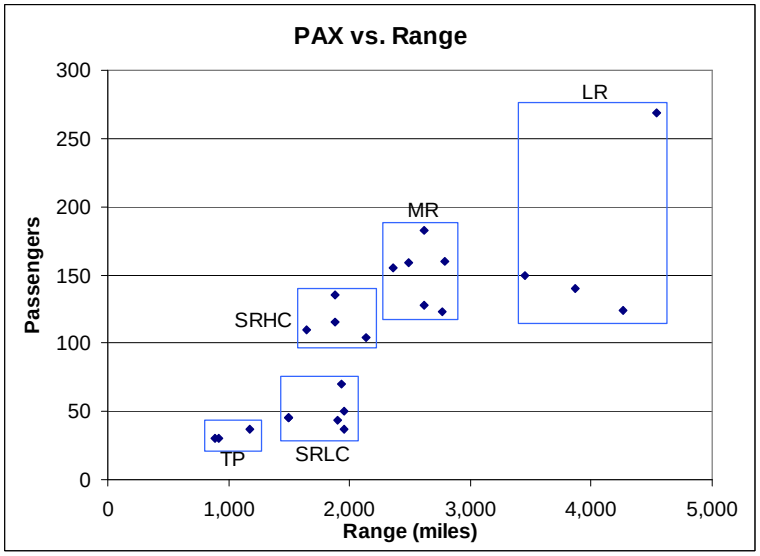


Figure 66: Selecting five aircraft types

and 1960. In 1966, the ATA published a formula for operating cost for the last time with the document title "Standard Method of Estimating Direct Operating Cost," which is the bases of today's practice of estimating DOC (Radnoti, 2002, chap. 3). A data-driven approach was used in this research. A good database for assessing VC is air carrier financial data of 'Schedule P-52' from US DOT From 41, which has DOC reported quarterly by fleet type for large certificated U.S. air carriers. This data includes depreciation as part of DOC. Since depreciation of the capital value of an aircraft can be widely different due to individual airline's situation and depreciation period, it is decided to separate it from the flight-related costs in the variable cost assessment in this research.

Fixed cost can be assessed from the airplane price. But airplane prices are tricky numbers since most of the airlines get special deals and do not pay the list prices. Therefore, assessed lease price was used for fixed cost. Table 15 shows the subjectively estimated lease prices for the representative airplane classes. Year 2005 was selected because Schedule P-52 database started listing 'Total Aircraft Airborne Hours' since year 2003 and also it is better not to use the data collected at the early stages of the

practice since many database tend to have errors when a new practice starts.

Table 15: Five selected aircraft classes

Class	PAX	Range (mile)	Speed (mph)	DOC / AirHr	Lease/Mo Est.($1,000)	Description
TP	32	991	316	1230	80	Turbo Prop
SRLC	49	1724	526	1778	100	Short Range Low Capacity
SRHC	114	1816	556	3656	150	Short Range High Capacity
MR	150	2569	503	3519	200	Medium Range
LR	145	3872	516	2798	260	Long Range

4.2.4.2 Airborne time calculation

Airplane segmental flight time is calculated using the method outlined here. Airborne time of a flight can be divided into ascending, cruise, and descending phase at its simplest form. Figure 67 shows schematic view of these flight portions. In the figure, t_A represents ascent, t_C for cruise, and t_D for descent. Constant speed is assumed during cruise.

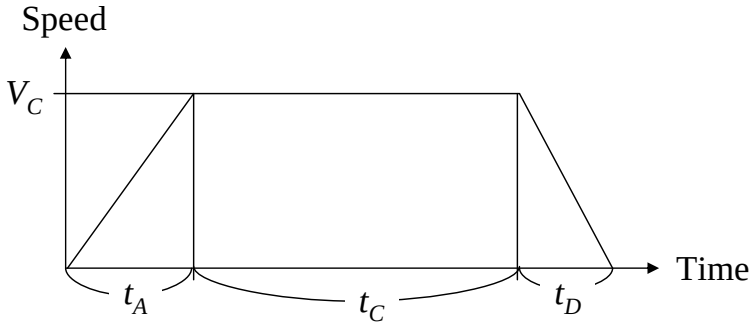


Figure 67: Simplified flight profile

Lewe (2005) suggests that t_A and t_D needs to be accounted for to increase accuracy of a simulation. The following equation suggested by Lewe is used to assess t_A and t_D in this research.

$$t_{A,C,D} = \frac{D_R}{V_C} + T_{A,D} \tag{12}$$

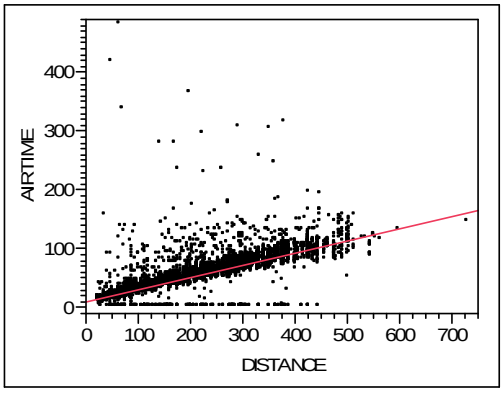
where D_R, V_C, and $T_{A,D}$ represent segment distance, cruise speed, and sum of ascent and descent time ($t_{A,D} = t_A + t_D$) respectively.

To determine t_A and t_D, a database study on T100D_SEGMENT was conducted. As in other database on NAS, T100D is also error prone. Figure 68a drawn for turbo prop (TP) class shows many data that have almost zero or irrationally high airtime for their trip distance. After removing the data points that took less than the flight time of cruise phase only without t_A or t_D (i.e. D_R/V_C) and that took more than two hours for $T_{A,D}$ to eliminate outliers, new graph was drawn as shown in Figure 68b. This elimination process was performed using the equation below,

$$D_R/V_C < AirTime < D_R/V_C + 120$$

where the units for time and distance are in minutes and miles.

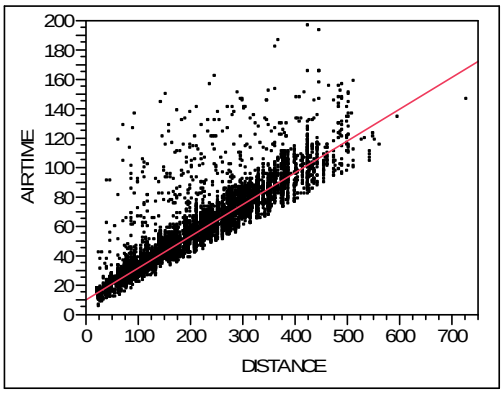
The reference cruise speed (V_C) used for the equation is 'Speed' column in Table 15. Representative $T_{A,D}$ is obtained as the y-intersect and the operational cruise speed $V_{C,OP}$, not the reference cruise speed by aircraft manufacturer (V_C), is obtained by using the relationship of $V_{C,OP} = 1/slope$. $T_{A,D}$ for other classes of the airplanes are obtained in the same way. Figure 69 shows the data plot.



AIRTIME = 10.758023 + 0.2064101*DISTANCE

R^2 = 0.58814

(a) Data fit with erroneous entries



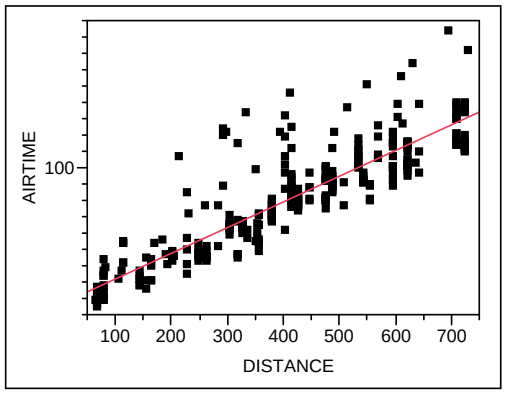
AIRTIME = 10.234961 + 0.216452*DISTANCE

R^2 = 0.821053

(b) Data fit without outliers

Figure 68: Finding $T_{A,D}$ for turboprop class

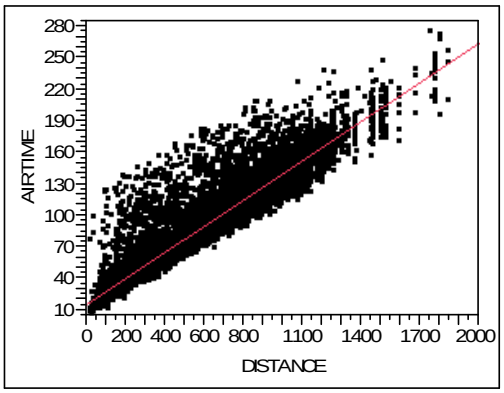
The results for all the five aircraft types have been determined by regressing the data points and setting $V_{C,OP}$ to $1/slope$ and $T_{A,D}$ to y-intersect. The results are



AIRTIME = 15.930685 + 0.1569154*DISTANCE

R^2 = 0.7745

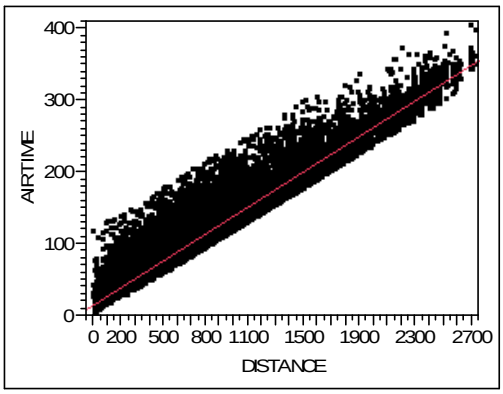
(a) Data fit for SRLC



AIRTIME = 15.549636 + 0.1244541*DISTANCE

R^2 = 0.915708

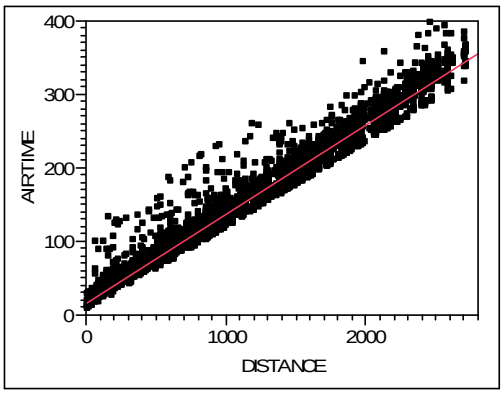
(b) Data fit for SRHC



AIRTIME = 14.374159 + 0.1248655*DISTANCE

R^2 = 0.957949

(c) Data fit for MR



AIRTIME = 15.145148 + 0.1214291*DISTANCE

R^2 = 0.97422

(d) Data fit for LR

Figure 69: Finding $T_{A,D}$ for SRLC, SRHC, MR, LR Class

listed in Table 16. One thing to note is that the speed of SRLC shows big disparity from the published cruise speed. It tells us that published data may not represent real operational speed when the distance is too short to have optimal cruise speed.

Table 16: Operational speed and ascent/descent time for five representative classes

Class	$V_{C,OP}$	t_{AD}	Description
TP	277	10.2	Turbo Prop
SRLC	382	15.9	Short Range Low Capacity
SRHC	482	15.5	Short Range High Capacity
MR	481	14.4	Medium Range
LR	494	15.1	Long Range

Another approach can be taken when the operational cruise speed has to be fixed to a certain value. In this case, the slope is pre-determined as $slope = 1/V_{C,OP}$ and $T_{A,D}$ becomes y-intersect of a regressed line with the slope that generates the least error with respect to the regression. Since the speed in Table 15 are not from the actual operational data but from published ones, it is agreed that data from T100D represents the real world better.

To calculate the time spent to travel from airport i to airport j, the distance needs to be known. In this simulation, distance is calculated as the great-circle distance as an approximation. Great circle distance is the shortest path between two points on the surface of a sphere. The center of a great circle coincides with the center of the sphere. The two points divides the great circle into two arcs and the shorter arc is the great circle distance. The earth is approximately spherical with the radius ranging from 3,949.901 (polar) to 3,963.189 (equatorial) miles. In the calculation, authalic mean radius of 3,958.760 mile was selected as the mean radius of the earth. Geographical formula for great circle distance is as follows. The distance (d) on a sphere is the product of radius (r) and angular difference ($\Delta\hat{\sigma}$), e.g. $d = r \times \Delta\hat{\sigma}$. Angular difference ($\Delta\hat{\sigma}$) can be represented using the spherical law of cosines as : $\Delta\hat{\sigma} = arccos(cos\phi_1 cos\phi_2 cos\Delta\lambda + sin\phi_1 sin\phi_2)$, where 1 and 2 refers to each of the

two points and ϕ, λ, Δ for latitude, longitude, and difference respectively. Since this equation can have large rounding off errors when distance are small, haversine formula was used for the calculation. Using this formula, the angular difference is represented as: $\Delta\hat{\sigma} = 2arcsin(\sqrt{sin^2(\frac{\phi_2 - \phi_1}{2}) + cos\phi_1 cos\phi_2 sin^2(\frac{\Delta\lambda}{2})})$

4.2.4.3 Viability consideration

Airlines do not provide service if they cannot generate economic or strategic benefits from it. However, detailed modeling of their behavior is beyond the scope of this thesis. Therefore, two simple considerations were made to prevent the simulation from providing a service when demand is below a threshold, as described below.

1. Cell operation will not be performed for an O-D pair whose demand cannot fill the half of the smallest airplane. (Load factor of 0.5 for the smallest aircraft is the criteria for viability consideration.)

2. After the cell operation, if there is a routing that does not meet the same criteria, it is distributed on the viable routes proportionally.

Figure 70 illustrates this process. In this example, the first segment of the upper route is assumed that it does not meet the viability criteria (dotted line). Therefore, the previously assigned demand of size 10 is distributed among the two viable routes proportionally. The remainder from the calculation is given to the viable routes in descending order.

4.2.5 Integration of components

Figure 71 shows the relationships between classes of the simulation framework. With the O-D demand (τ) matrix to be served, traveler objects are generated with the demand information and stored in the object of the origin airport. Airline object retrieve the demand information by accessing each airport object and collecting information on the travelers at the airport. Airlines give ticket object to traveler objects

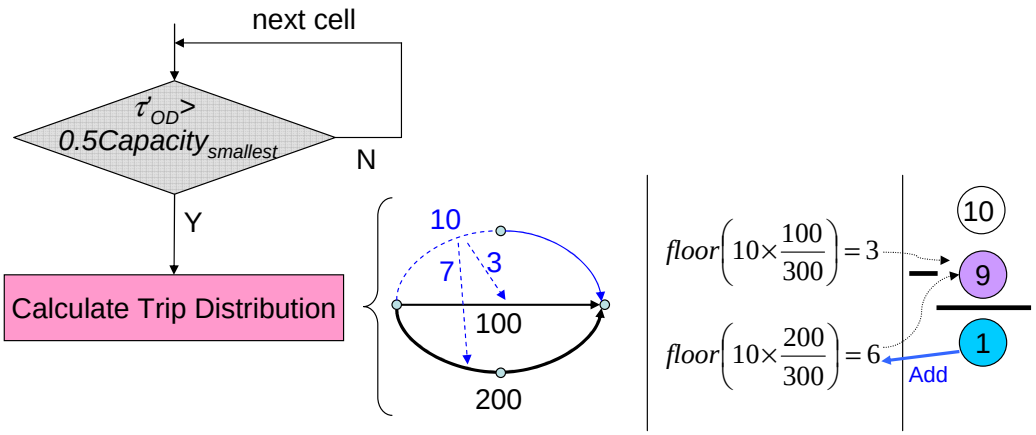


Figure 70: Operational viability consideration

for different routes according to the demand distribution from the probabilistic choice model. Aircraft type and number of aircrafts and operations are determined on the segment pairs that are part of the routes to serve the changed segmental demands. In allocating the airplanes, they are added to the higher ranked airport first and then to lower ranked one on each segment pair and take turns. This way each airport on each segment pair has equivalent number of aircrafts with the same or one more aircraft on the higher ranked airport. When the trip distribution process ends, architect object adds new airports to the network and incremental demand according to the next evolution point. The process repeats until the calculation for the last evolution point is finished.

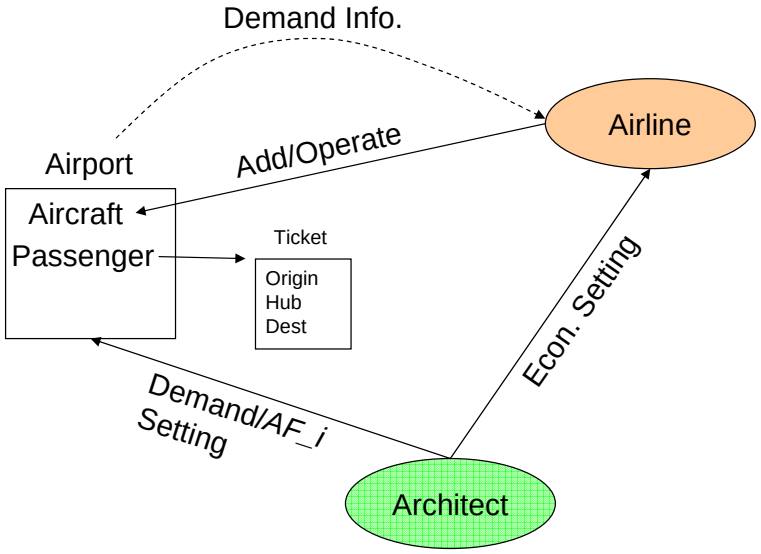


Figure 71: Class relationship diagram

CHAPTER V

SIMULATION

5.1 Theoretical baseline determination

In this section, the process that was taken to determine theoretical baseline is described. Theoretical baseline is the one that captures the trends observed in the real NAS when attraction factors (AF_i) are all neutral.

5.1.1 Selection of network parameters

As the case of the demand model, 66 large and medium hubs in year 1997 were selected, where 13 airports were categorized into 8 Metroplex airports. Therefore, 53 network nodes were used for the network model. The coordinates for the metroplex airports' centers were decided to be at the geometric centers of their component airports. The time span of the network simulation is one day. Therefore, the annual demand from sDB1B needs to be converted into daily demand. The following scale factor was used to convert the annual demand to daily one.

$$\text{dSF} = \text{totalEnp}_{\text{T100D}}/\text{totalEnp}_{\text{sDB1B}} \times \frac{1}{365}$$

where dSF: scale factor to change annual demand to daily demand.

$\text{totalEnp}_{\text{T100D}}$: total enplanements from T100D_SEGMENT data for the selected 66 airports.

$\text{totalEnp}_{\text{sDB1B}}$: total enplanements from sDB1B data from the travel between the selected 66 airports.

The values of dSF of each year were calculated with the above equation from year 1993 to 2005, which are given in Table 17.

Before running the active design framework, some network parameter values need

Table 17: Scale factor for conversion of yearly to daily demand

Year	$\text{totalEnp}_{\text{T100D}}$	$\text{totalEnp}_{\text{sDB1B}}$	dSF
1993	398,826,893	13,977,440	0.07817
1994	435,929,796	16,035,790	0.07448
1995	449,761,715	16,948,944	0.07270
1996	477,546,981	18,555,680	0.07051
1997	492,798,875	19,529,276	0.06913
1998	502,950,777	21,717,966	0.06345
1999	523,590,204	24,709,592	0.05805
2000	546,393,749	26,091,234	0.05737
2001	509,824,732	24,824,392	0.05627
2002	498,257,022	23,419,176	0.05829
2003	518,905,664	23,274,622	0.06108
2004	558,466,445	25,023,544	0.06114
2005	580,813,380	26,060,230	0.06106

to be determined. Table 18 shows the list of network parameters and the values deemed reasonable for the theoretical baseline determination.

Table 18: Parameters in the network model

Variables	Range	Base	Description
AF_i	$[-\infty\ ,\ \infty]$	0	A attraction factor for hub connection (0: neutral, positive: benefit, negative: penalty for hub connection)
limitLF	[0 , 1]	0.7	Load factor limit for aircraft
r_{qp}	[0 , 1]	0.1	Relative radius for determination of quasi-Pareto options
w	[0 , 1]		Weight factor for disutility calculation
α	$[0\ ,\ \infty]$		Scale factor used to make the difference of disutility more or less discernable (The higher, the more discernable)

w and α are the most influential parameters. Therefore, their effect was tested to determine the theoretical baseline of the framework. Five cases were selected as in Table 19.

The order of the airports that are added to the simulation is determined by the size of the enplanements of the target year. For generating the airport order for this study, the list the 53 airports ranked by the total enplanements from sDB1B in year

Table 19: Testing w and α for baseline determination

		a		
		1	5	25
w	0		X	
	0.5	X	X	X
	1		X	

2000 in descending order are selected. Table 20 lists the airports with their rankings.

Table 20: Airports ranked w.r.t. enplanements from sDB1B in year 2000

Rank	Airport	Enp.	Rank	Airport	Enp.	Rank	Airport	Enp.
1	LA_	99,264	19	PHL	25,981	37	MEM	10,526
2	CH_	98,679	20	SAN	25,433	38	CMH	10,427
3	NY_	94,031	21	TPA	23,049	39	MKE	10,205
4	SF_	83,311	22	SLC	21,832	40	ABQ	9,802
5	DA_	77,505	23	MCI	19,460	41	PBI	9,468
6	ATL	69,808	24	PDX	17,489	42	RNO	9,071
7	WA_	66,264	25	CLE	17,480	43	PVD	8,617
8	PHX	58,293	26	CLT	16,582	44	JAX	7,932
9	LAS	54,174	27	CVG	15,962	45	RSW	7,585
10	DEN	50,798	28	RDU	15,891	46	OMA	6,731
11	HU_	50,416	29	PIT	15,842	47	BUF	6,162
12	STL	45,791	30	SMF	15,173	48	SDF	6,120
13	MCO	40,341	31	MSY	14,694	49	TUS	5,810
14	MI_	38,355	32	BNA	13,238	50	TUL	5,728
15	MSP	36,416	33	IND	11,561	51	OKC	5,542
16	DTW	35,566	34	BDL	11,483	52	ELP	5,112
17	BOS	35,280	35	AUS	11,193	53	COS	4,104
18	SEA	29,737	36	SAT	10,578			

5.1.2 Sensitivity study

Figure 72 shows O-D demand (τ) and enplanements comparison of sDB1B of year 2000. The difference between them comes from the enplanements boosts by connections.

After setting the parameter values and determined the order of the airport entries, several evolution paths were tested to generate calibrated results against the data from sDB1B of year 2000. Figure 73 shows the three paths that were taken for the

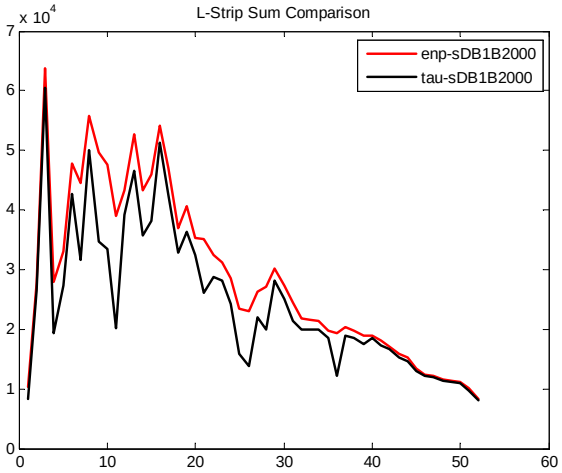


Figure 72: Comparison of τ and enplanements matrix of year 2000

analysis. The results from each evolution path is given in the following subsections one by one. The the attraction factors for connecting flights were neutral ($AF_i = 0$). Preferential attachment is indirectly applied because the airports are ordered with the high demand first and the network is constructed connecting to these core airports first, utilizing the economies of scale, thus providing attractiveness naturally. The distance demand histograms in the following sections have the bin size of 50 miles. Affinity measures used for the study was based on CDF of each l-strip. This gives penalty when numbers do not match for highly demanded airports which consists majority of trips. Affinity measure for distance demand was based on CDF of each bin. This also gives penalty when short-range demand is not represented correctly since the majority of distance demand is short-range trips.

5.1.2.1 Path 1: Spatial-only expansion

The first path chosen for the test was spatial only expansion. In this scheme, each airport is added to the simulation with its full demand. Figure 74, 75, 76, 77, and 78 show the five cases where $(w = 0, \alpha = 5)$, $(w = 0.5, \alpha = 1)$, $(w = 0.5, \alpha = 5)$, $(w = 0.5, \alpha = 25)$, and $(w = 1, \alpha = 5)$ respectively on path 1.

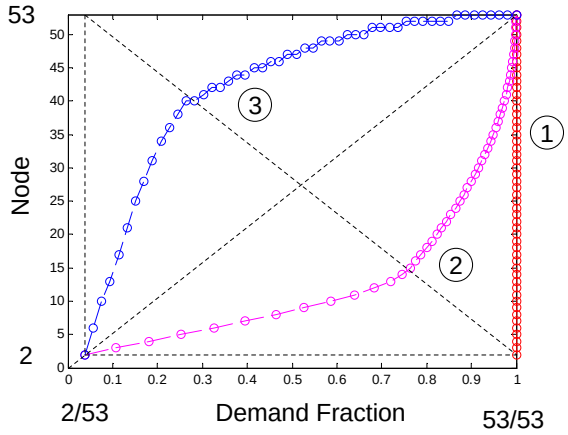
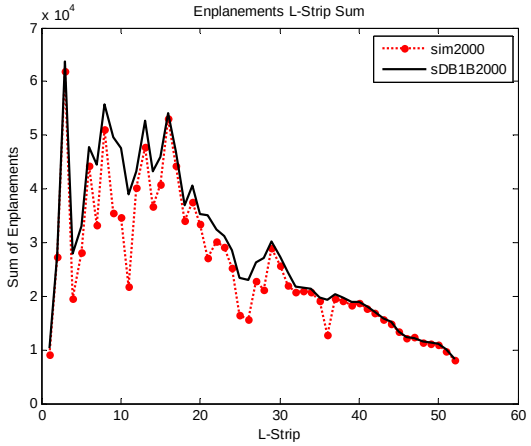
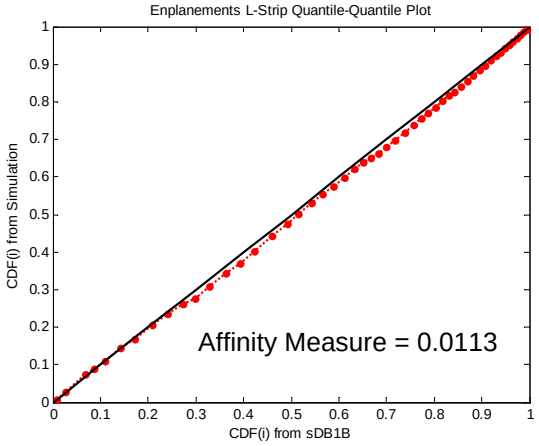


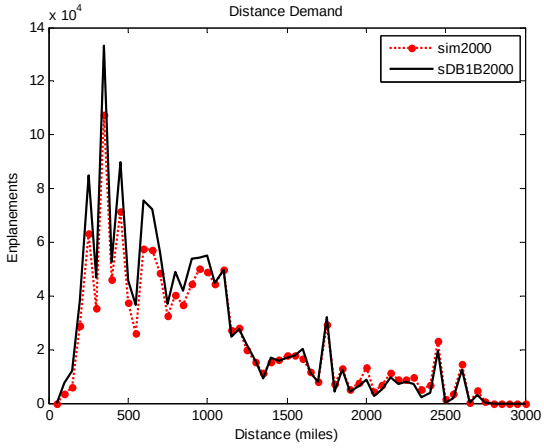
Figure 73: Evolution paths taken to select baseline



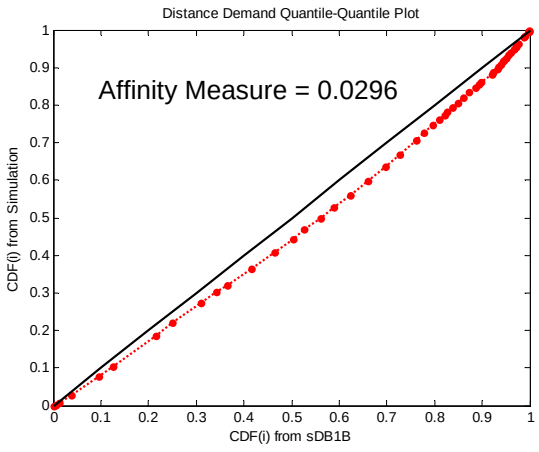
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

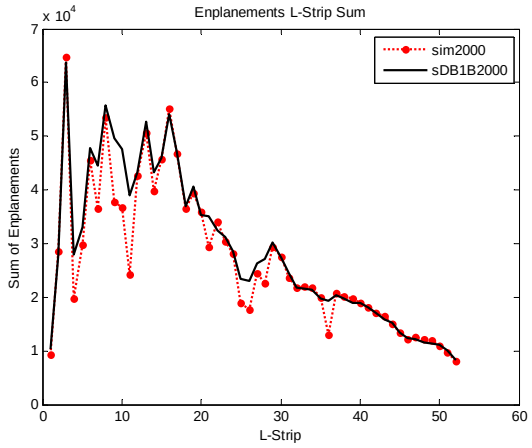


(c) Distance demand

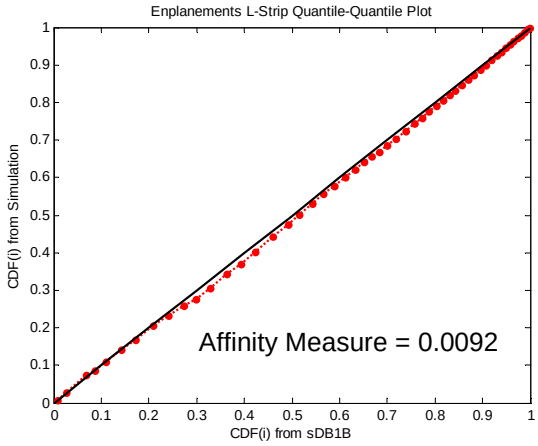


(d) Distance demand QQ plot

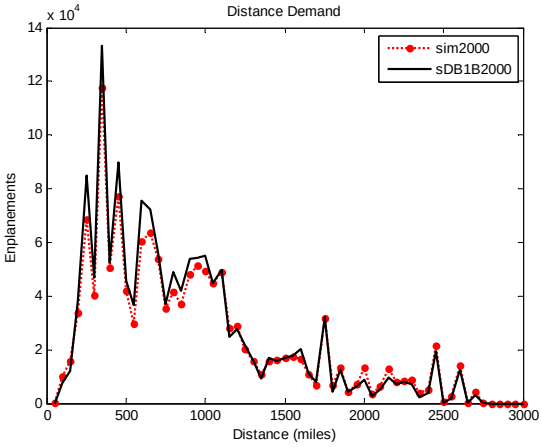
Figure 74: Results: $w = 0$ and $\alpha = 5$ on spatial-only expansion



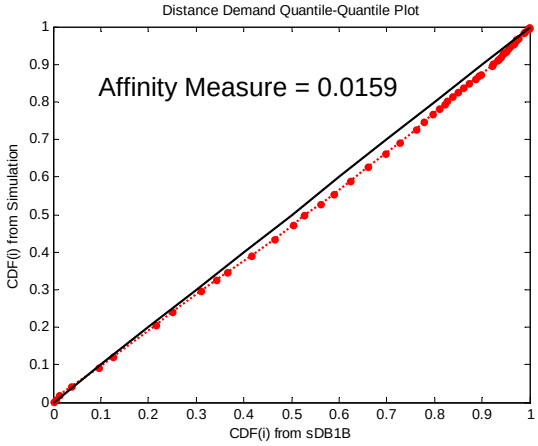
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

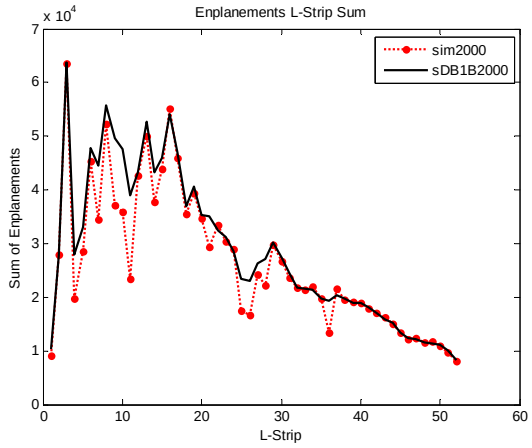


(c) Distance demand

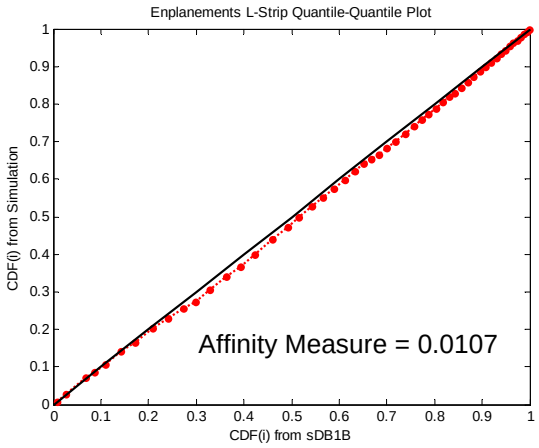


(d) Distance demand QQ plot

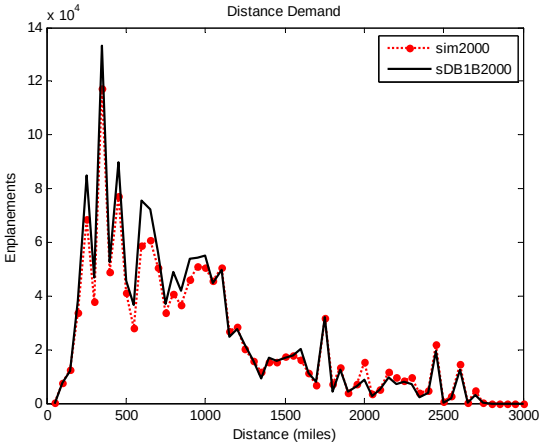
Figure 75: Results: $w = 0.5$ and $\alpha = 1$ on spatial-only expansion



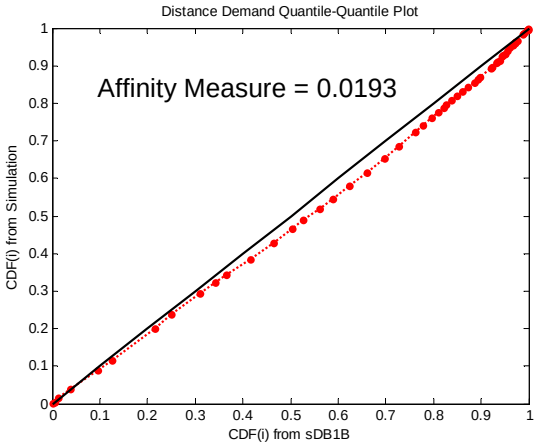
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

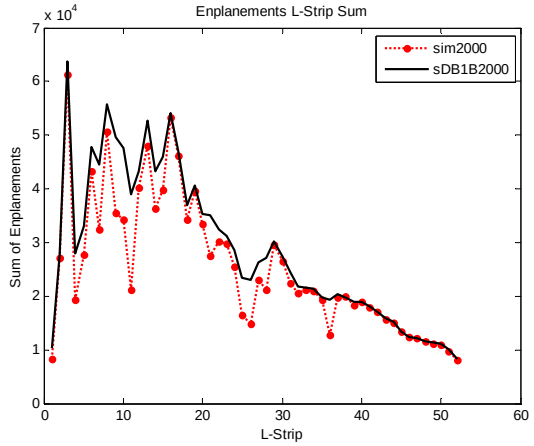


(c) Distance demand

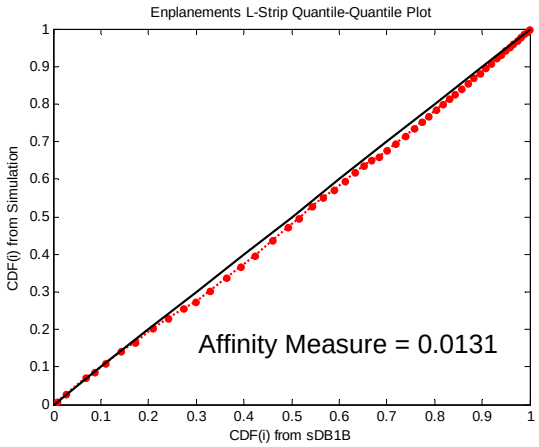


(d) Distance demand QQ plot

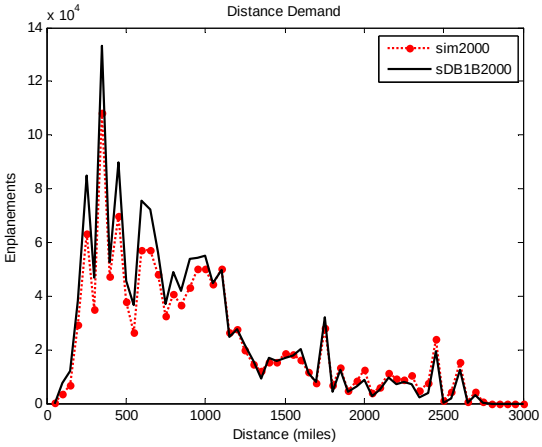
Figure 76: Results: $w = 0.5$ and $\alpha = 5$ on spatial-only expansion



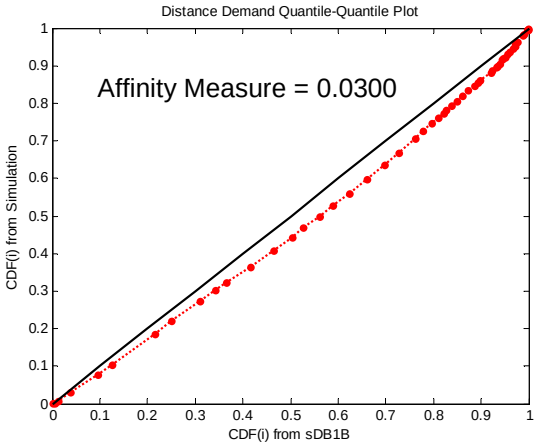
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

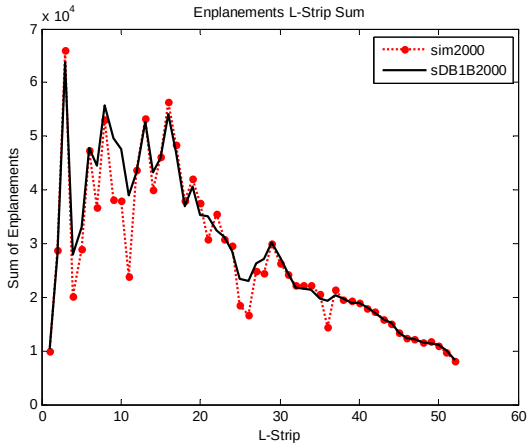


(c) Distance demand

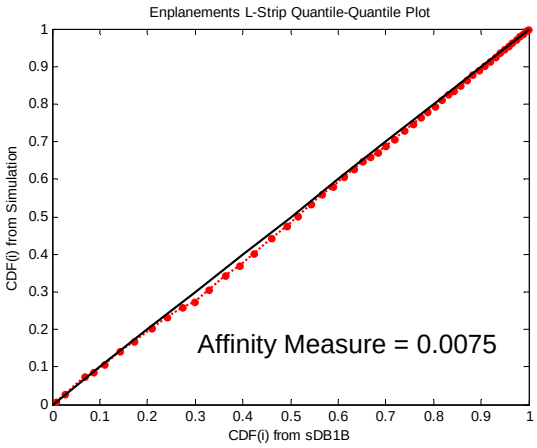


(d) Distance demand QQ plot

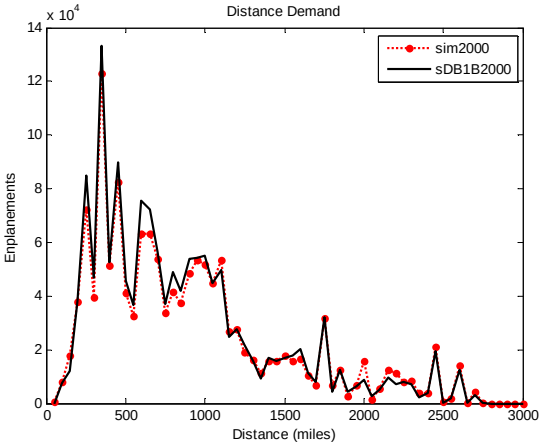
Figure 77: Results: $w = 0.5$ and $\alpha = 25$ on spatial-only expansion



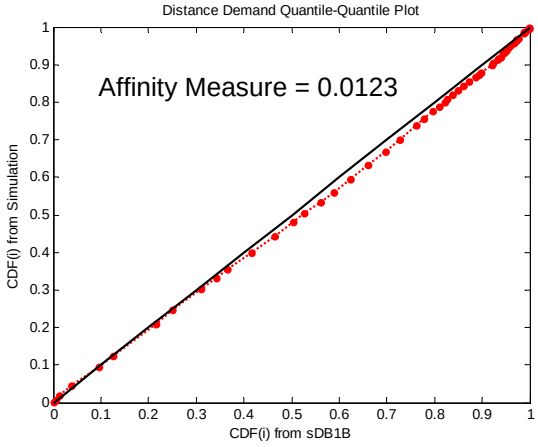
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot



(c) Distance demand

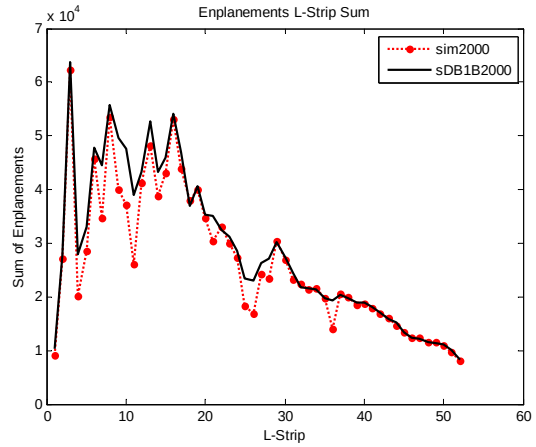


(d) Distance demand QQ plot

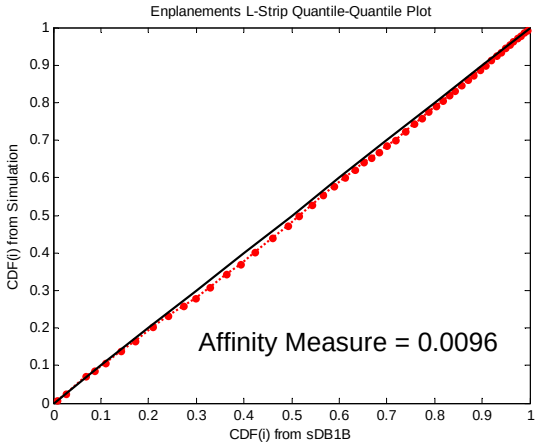
Figure 78: Results: $w = 1$ and $\alpha = 5$ on spatial-only expansion

5.1.2.2 Path 2: Convex path

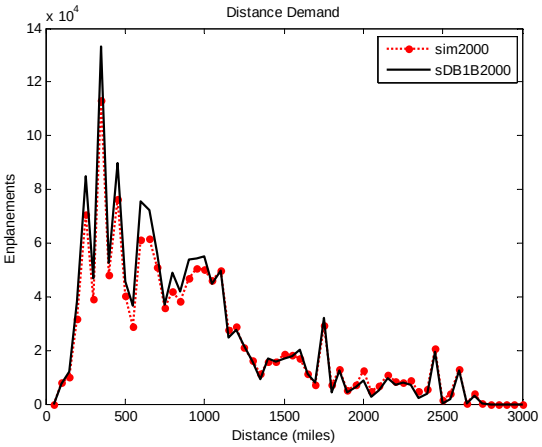
The evolution points selected for this analysis were generated with piecewise cubic Hermite interpolation function in Matlab with the three points of (0.0378,2), (0.7594, 14.7500), and (1, 53), among which the first and last are the start and end point of the evolution and the middle point reside on the 135 degree diagonal line at three quarter of its length from the left. Figure 79, 80, 81, 82, and 83 show the five cases where $(w = 0, \alpha = 5)$, $(w = 0.5, \alpha = 1)$, $(w = 0.5, \alpha = 5)$, $(w = 0.5, \alpha = 25)$, and $(w = 1, \alpha = 5)$ respectively on path 2.



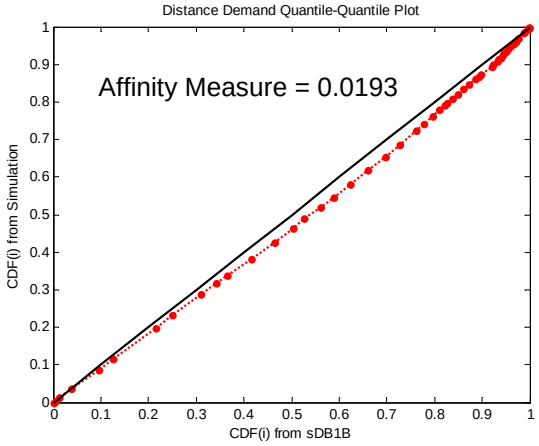
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

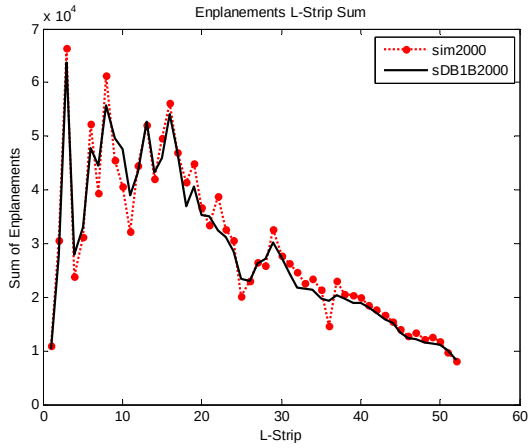


(c) Distance demand

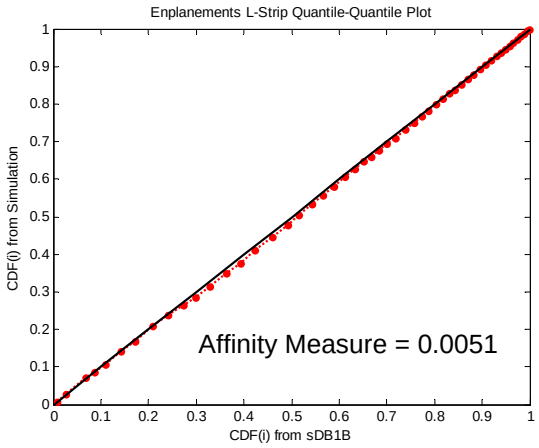


(d) Distance demand QQ plot

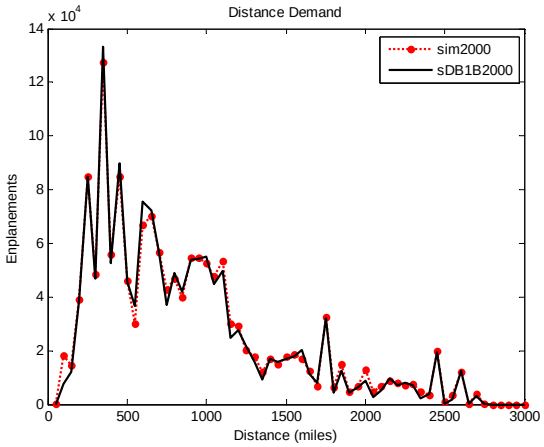
Figure 79: Results: $w = 0$ and $\alpha = 5$ on convex path



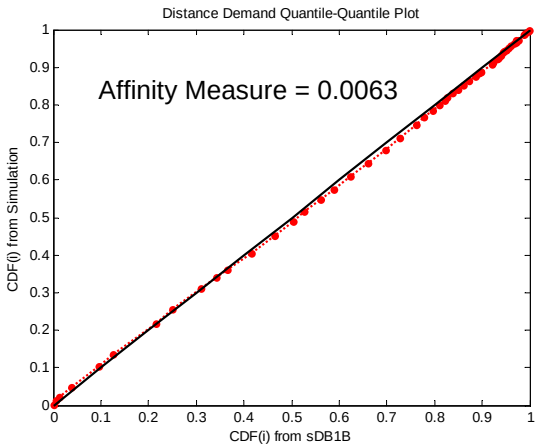
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

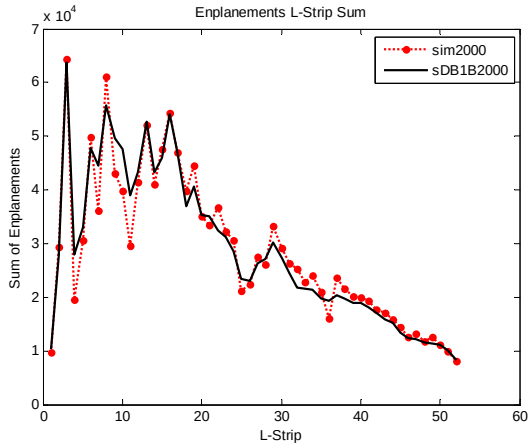


(c) Distance demand

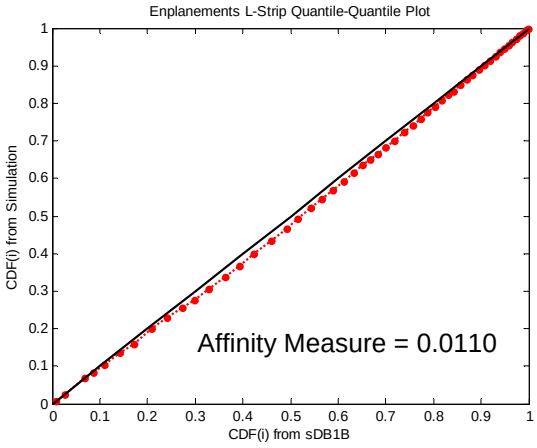


(d) Distance demand QQ plot

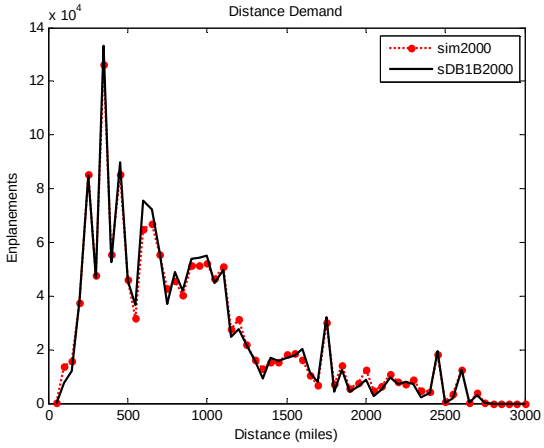
Figure 80: Results: $w = 0.5$ and $\alpha = 1$ on convex path



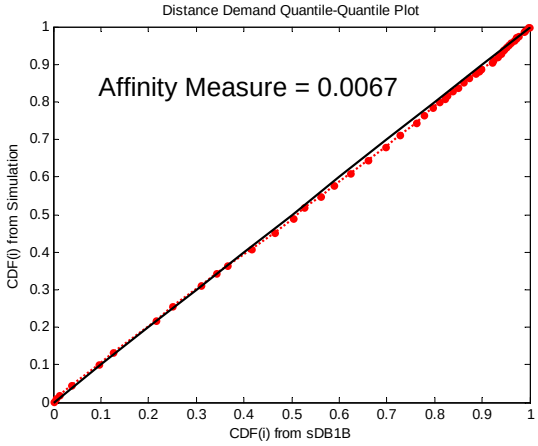
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

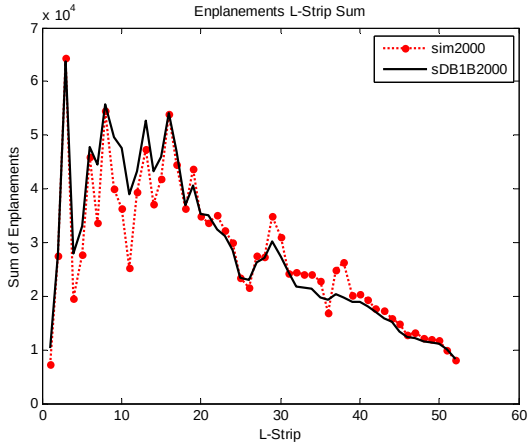


(c) Distance demand

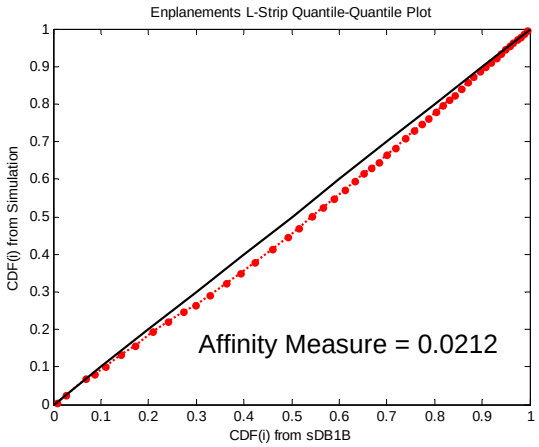


(d) Distance demand QQ plot

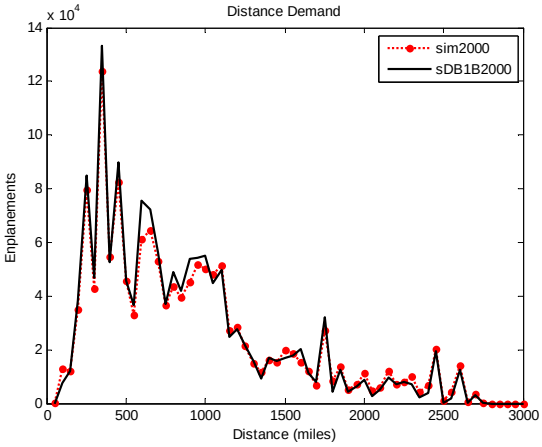
Figure 81: Results: $w = 0.5$ and $\alpha = 5$ on convex path



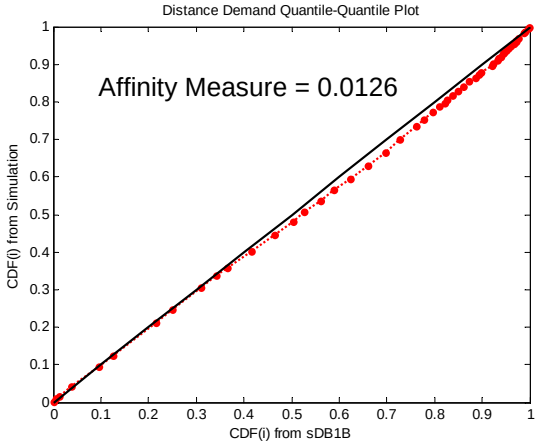
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

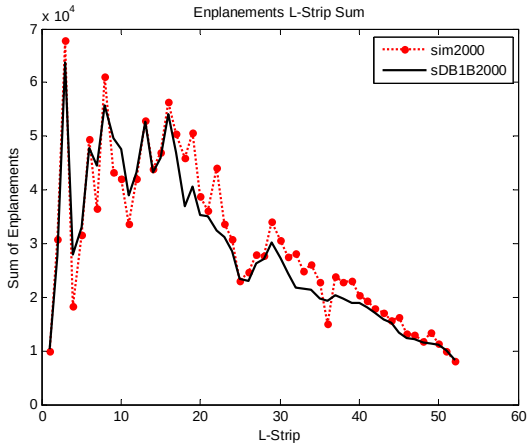


(c) Distance demand

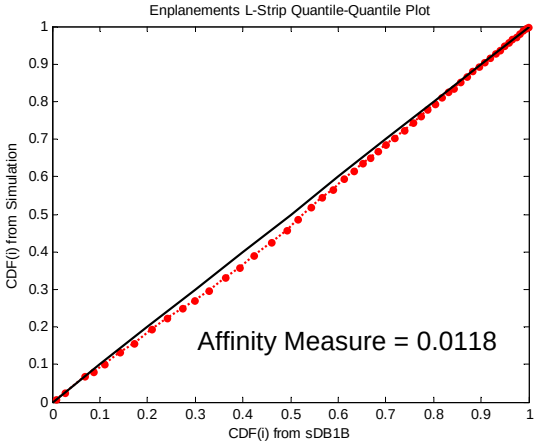


(d) Distance demand QQ plot

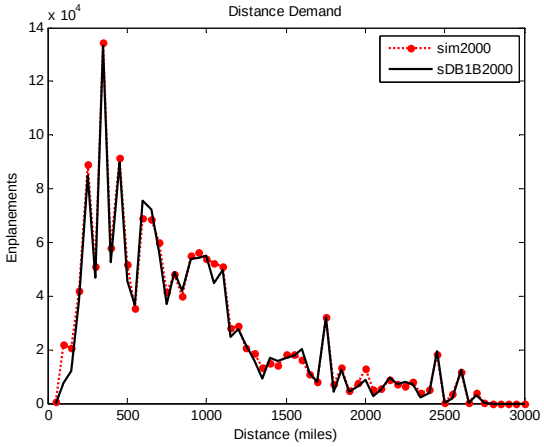
Figure 82: Results: $w = 0.5$ and $\alpha = 25$ on convex path



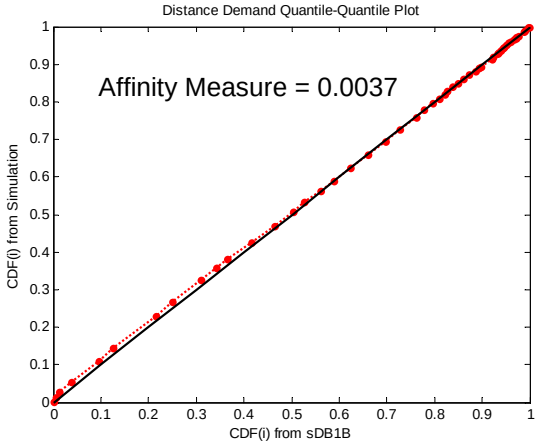
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot



(c) Distance demand

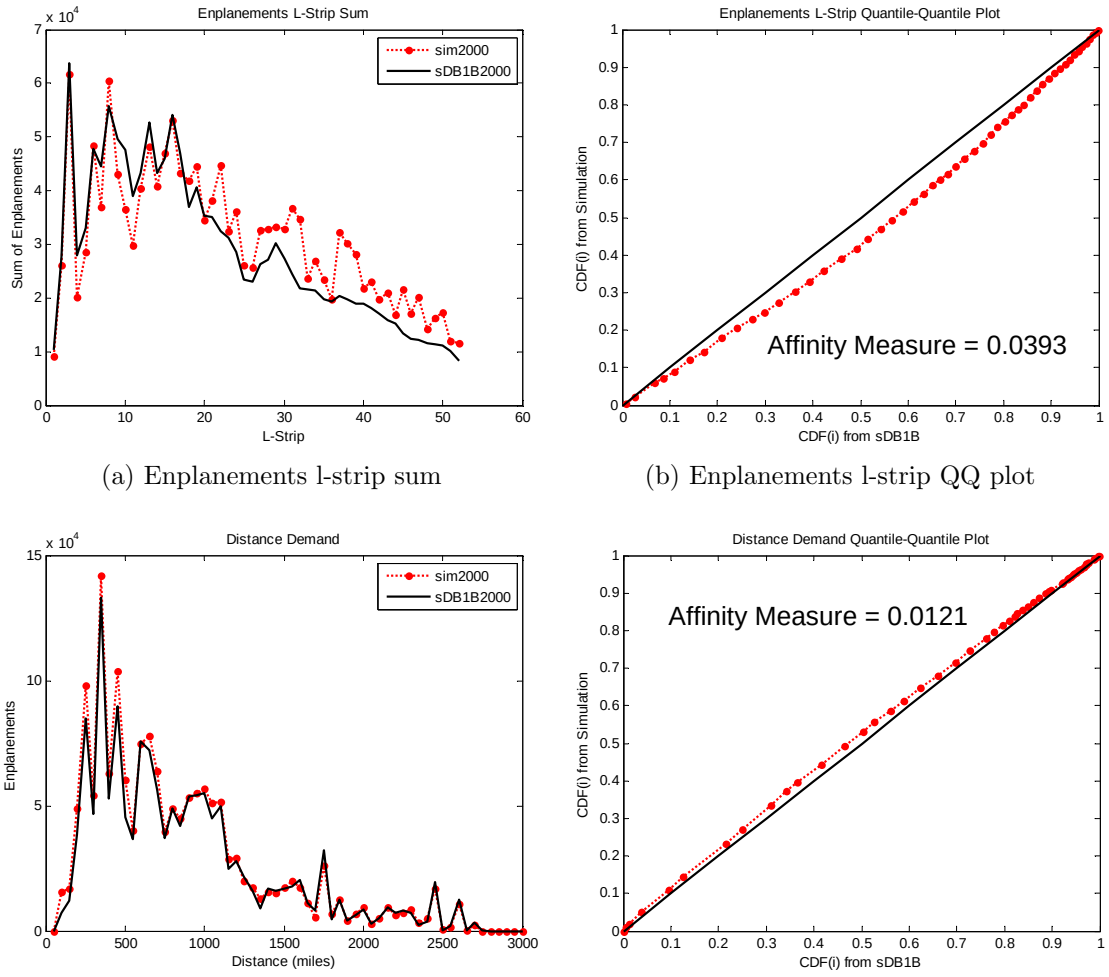


(d) Distance demand QQ plot

Figure 83: Results: $w = 1$ and $\alpha = 5$ on convex path

5.1.2.3 Path 3: Concave path

The evolution points selected for this analysis were generated by mirroring the points of the convex path case against the diagonal line. In other words, the concave and convex points are symmetric against the diagonal line. Figure 84, 85, 86, 87, and 88 show the five cases where $(w = 0, \alpha = 5)$, $(w = 0.5, \alpha = 1)$, $(w = 0.5, \alpha = 5)$, $(w = 0.5, \alpha = 25)$, and $(w = 1, \alpha = 5)$ respectively on path 4.



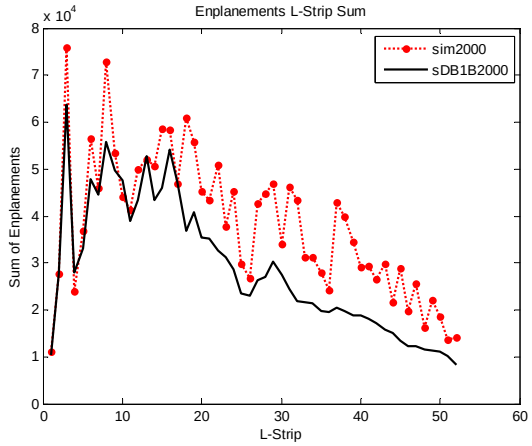
(a) Enplanements l-strip sum

(b) Enplanements l-strip QQ plot

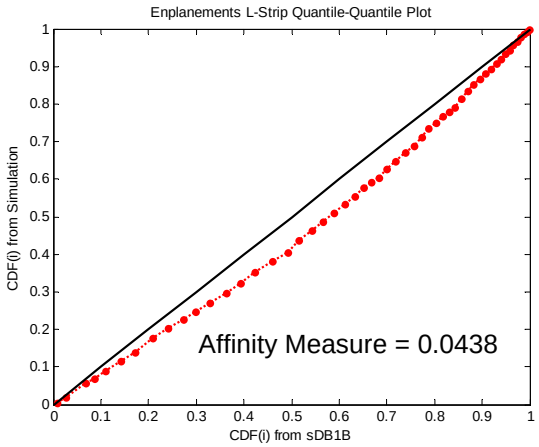
(c) Distance demand

(d) Distance demand QQ plot

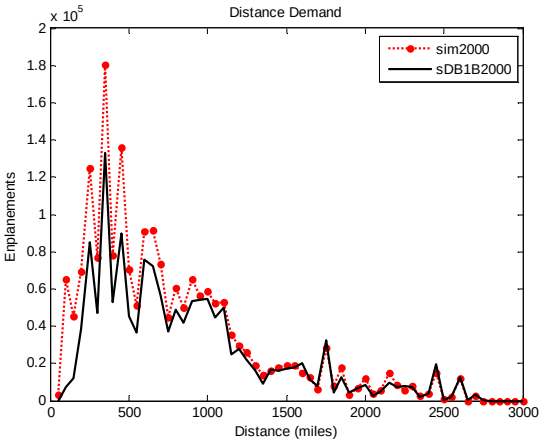
Figure 84: Results: $w = 0$ and $\alpha = 5$ on concave path



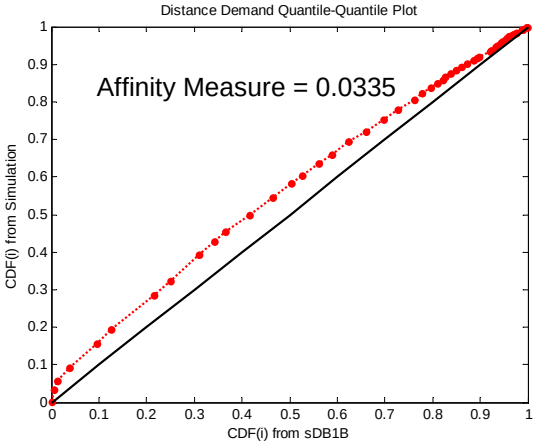
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

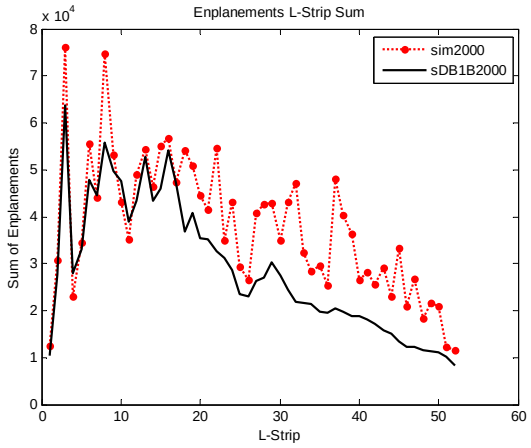


(c) Distance demand

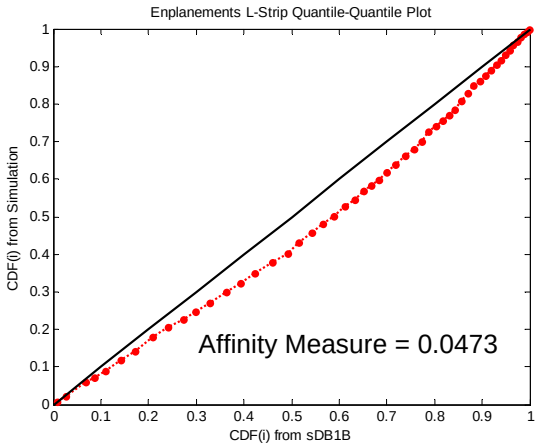


(d) Distance demand QQ plot

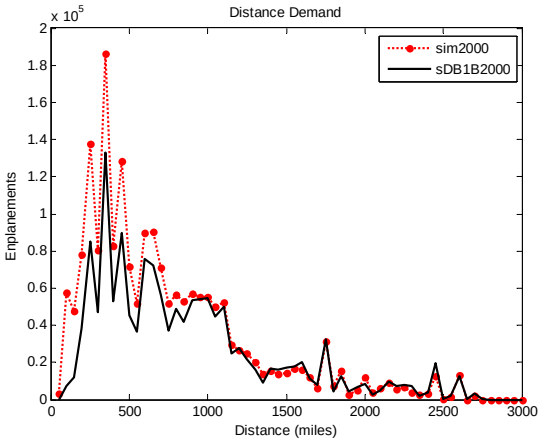
Figure 85: Results: $w = 0.5$ and $\alpha = 1$ on concave path



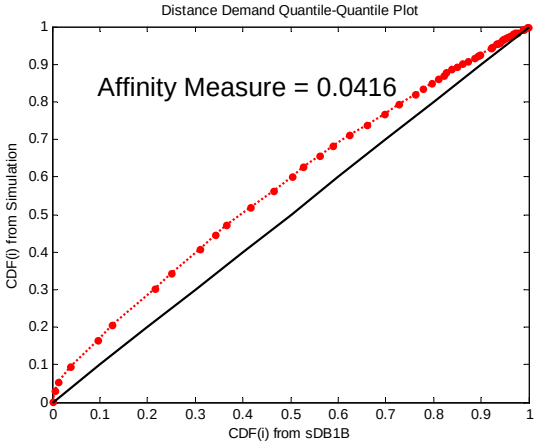
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

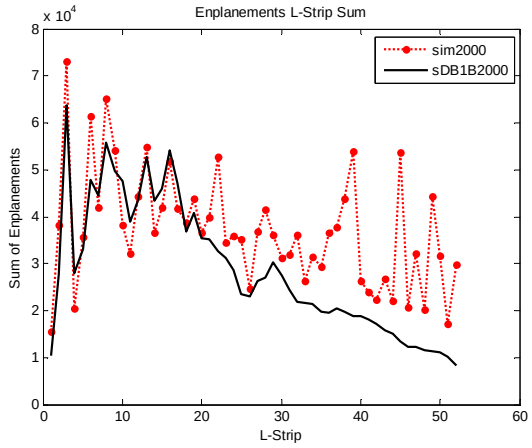


(c) Distance demand

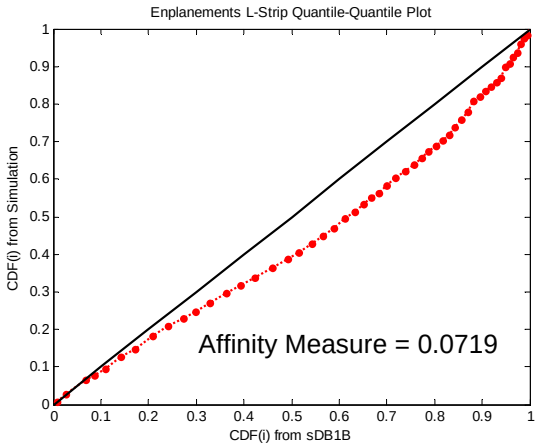


(d) Distance demand QQ plot

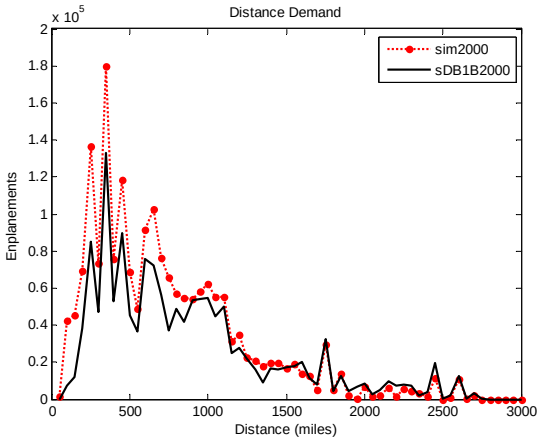
Figure 86: Results: $w = 0.5$ and $\alpha = 5$ on concave path



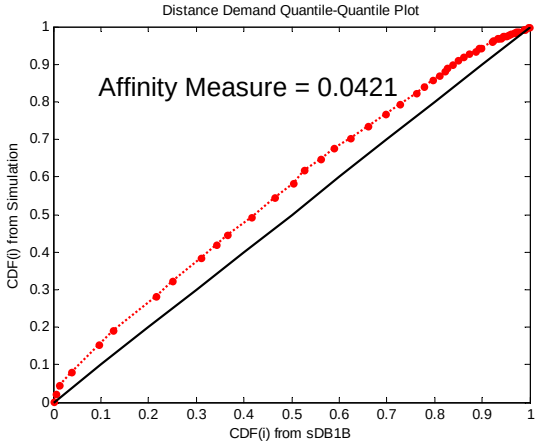
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot

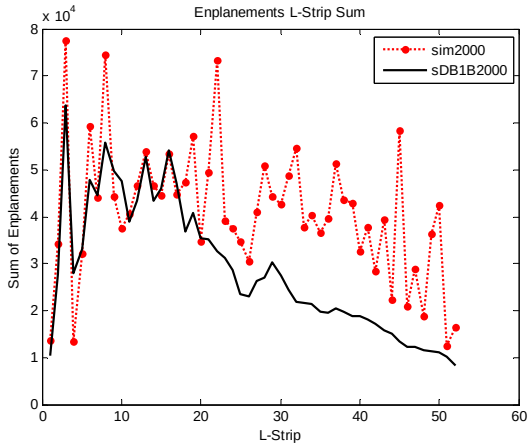


(c) Distance demand

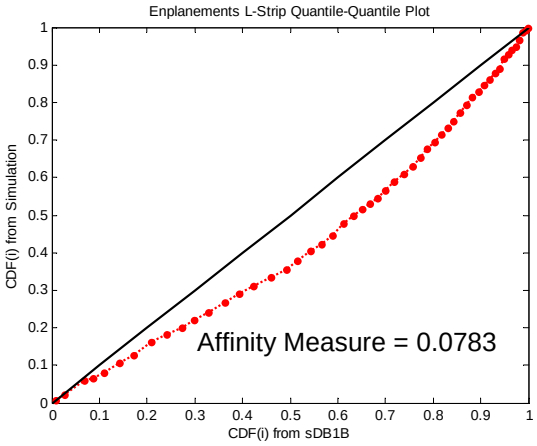


(d) Distance demand QQ plot

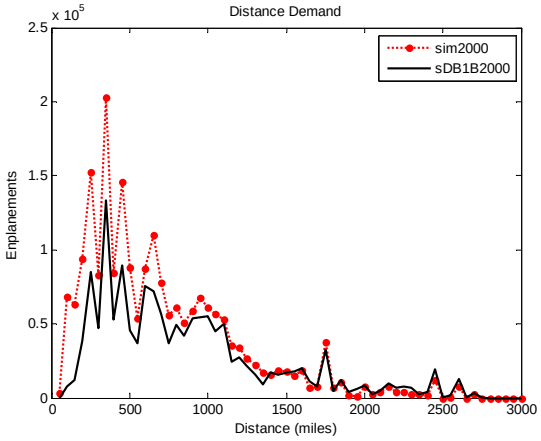
Figure 87: Results: $w = 0.5$ and $\alpha = 25$ on concave path



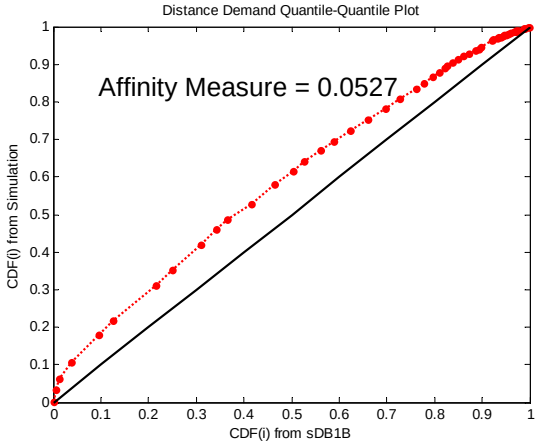
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot



(c) Distance demand



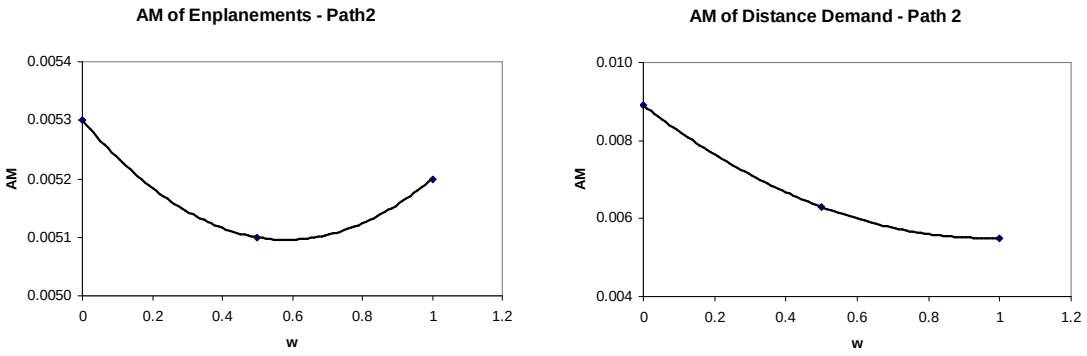
(d) Distance demand QQ plot

Figure 88: Results: $w = 1$ and $\alpha = 5$ on concave path

5.1.3 Baseline determination

The results from this study showed the best results when $\alpha = 1$ for all the three evolution paths. Therefore, the baseline value for α was set to 1. w values that are probable to generate better results (i.e. $w = 1$ for path 1, $w = 0$ and $w = 1$ for path 2, and $w = 0$ for path 3) were tested. A summary that shows Affinity Measure (AM) for enplanements l-strip sum and distance demand is in Table 21.

The results from the convex path progression (path 2) shows that it generated a network which is very close network to sDB1B. A diagonal progression that follows the 45 degree line was also tested but the convex path was superior. QQ plots of the convex path show lines that are most close to the diagonal line. The diagonal line on QQ plots describe the same distribution whether or not the magnitude are different. The best results with respect to AM for both enplanements and distance demand were also from the convex path. When second order polynomial graph was drawn for AM with respect to w when $\alpha = 1$ for enplanements as in Figure 89a, the value of w for the minimum is around $w = 0.6$. However, AM for distance is minimum at $w = 1$ as in Figure 89b.



(a) Affinity measures for enplanements (b) Affinity measures for distance demand

Figure 89: Affinity measures on convex path

Therefore, the baseline value for w was selected to be at 0.75 as a compromise. The results from the baseline is shown in Figure 90.

Table 21: Affinity measure values from baseline determination study

(a) Spatial: Enplanements

		a		
		1	5	25
	0		0.0113	
w	0.5	0.0092	0.0107	0.0131
	1	**0.0089**	0.0075	

(b) Spatial: Distance

		a		
		1	5	25
	0		0.0296	
w	0.5	0.0159	0.0193	0.0300
	1	**0.0134**	0.0123	

(c) Convex: Enplanements

		a		
		1	5	25
	0	**0.0053**	0.0096	
w	0.5	0.0051	0.011	0.0212
	1	**0.0052**	0.0118	

(d) Convex: Distance

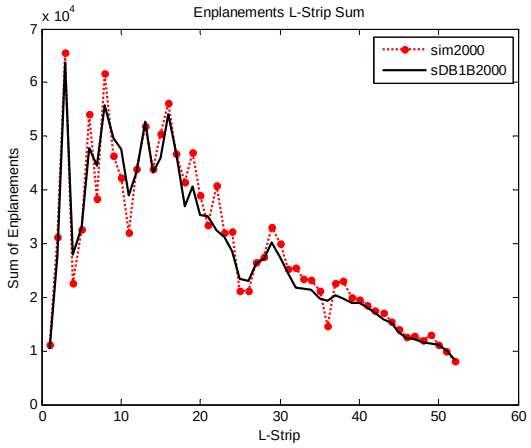
		a		
		1	5	25
	0	**0.0089**	0.0193	
w	0.5	0.0063	0.0067	0.0126
	1	**0.0055**	0.0037	

(e) Concave: Enplanements

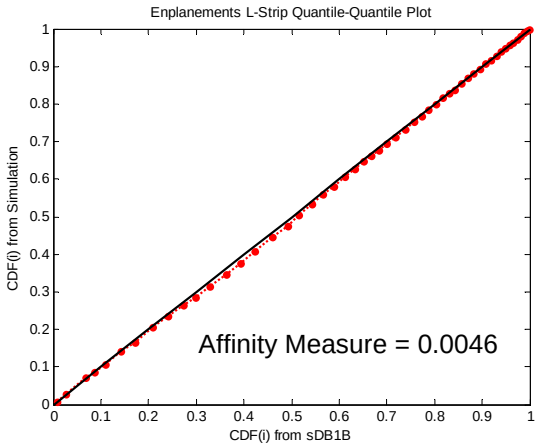
		a		
		1	5	25
	0	**0.0437**	0.0393	
w	0.5	0.0438	0.0473	0.0719
	1		0.0783	

(f) Concave: Distance

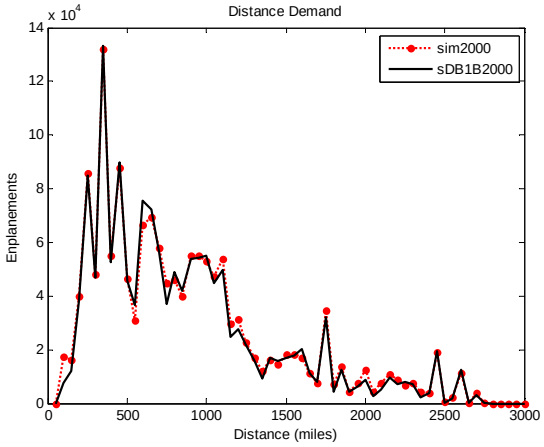
		a		
		1	5	25
	0	**0.0266**	0.0121	
w	0.5	0.0335	0.0416	0.0421
	1		0.0527	



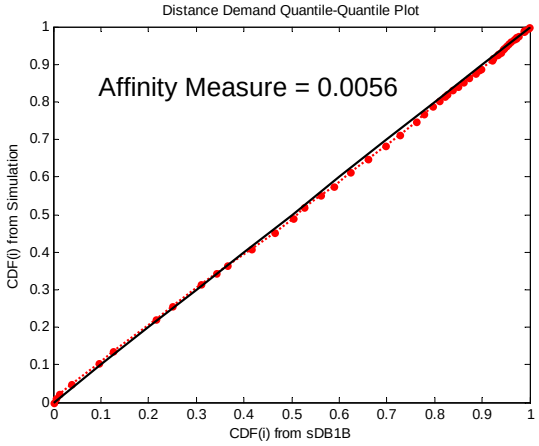
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot



(c) Distance demand



(d) Distance demand QQ plot

Figure 90: Results from baseline: $w = 0.75$ and $\alpha = 1$ on convex path

5.1.4 Findings

It was observed that the results from the active design algorithm is heavily dependent on the evolution path chosen. It can be visually observed that the total enplanements gets higher as the path deviates more from the spatial progression. It was also shown that it was possible to generate quite accurate results from the evolutionary network model by selecting a certain path without setting attraction factors. This reveals a significant finding regarding airline network modeling practice. It can be postulated that the chosen evolution path played a significant role in compensating for the incompleteness in the model to generate the desired result. It can be visualized with an example of making a mold of an object, depicted in Figure 91. In this example, we have a eccentrically shaped block object that needs to be molded. If we have a rigid mold that does not match the shape of the block, the cast will not be the same shape no matter how hard we try. But if we have a flexible container, whose shape can easily be changed, it is possible to generate well-fit mold by morphing it even when the initial mold had imperfections. The eccentrically shaped block is analogous to the true nature of the airline network and the mold is to a simulation that mimics it.

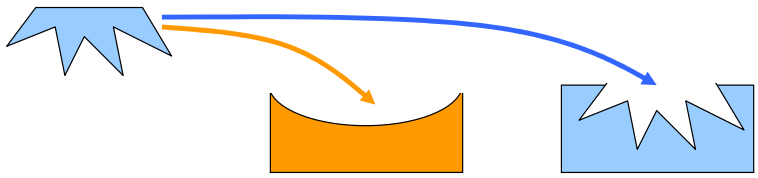


Figure 91: Analogy of airline network and a simulation to a eccentric block and a mold

It was also observed that the evolution path that resulted in a network similar to sDB1B has convex shape. This means that the considered network is best modeled in the accelerated evolution scheme if small number of airports are evolved first to a certain critical point to form the initial structure of the core network and the

rest of the airports are entered in a very short time period. This corresponds to the Matcalfe's law stating that network evolves slowly first and it speeds up after it reaches critical mass. Also, it should be noted that we are looking at a subset of the NAS. Therefore, it can be postulated that when all the airports in the NAS are considered, the proper evolution path might demonstrate an 'S' shaped curve extending the path obtained from this research. This postulation is made from the observation that many networked system sets the initial structure slowly, expand very fast until they mature, and then the expansion slows down as it approaches its full maturity. Figure 92 describes that the evolution path from this research might be a part of the overall picture.

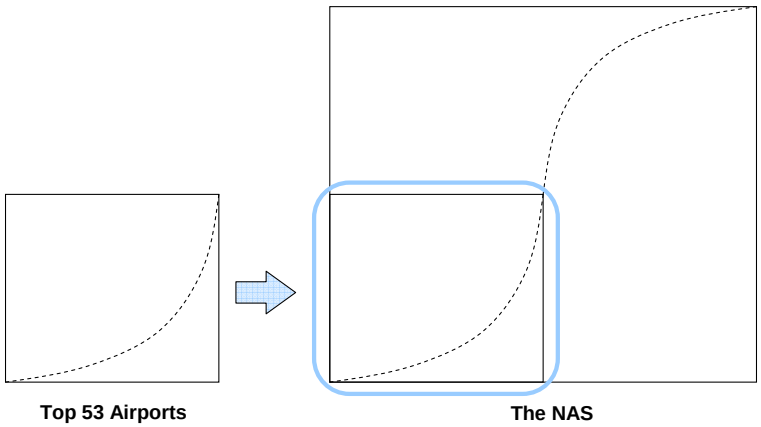


Figure 92: Possible shape of overall evolution path

5.2 Practical baseline determination

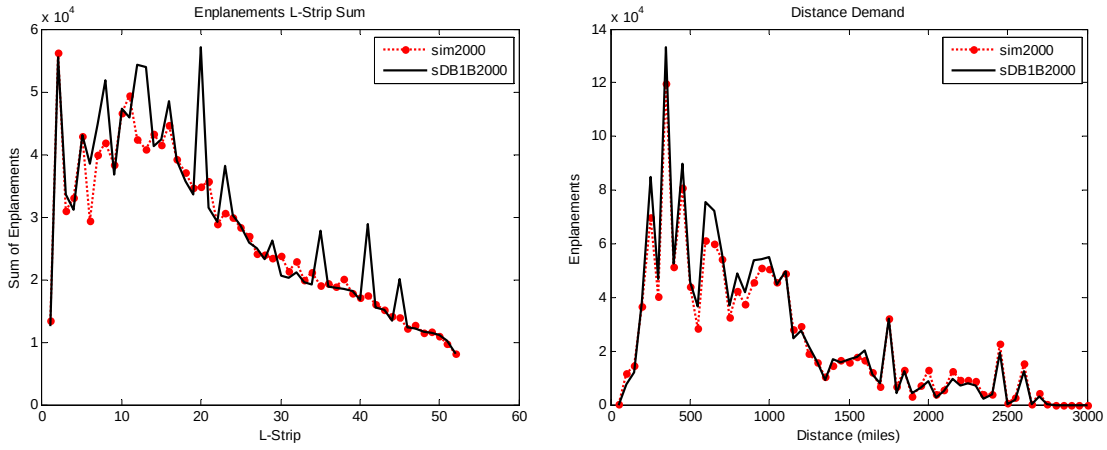
In this section, fine tuning of the model parameters was performed with the adjustments of attraction factors of individual airports to generate the practical baseline. The studies in this section were conducted using spatial progression on evolution space. Spatial progression was selected since the convex path takes significantly more time to follow — 20 hours vs. 30 minutes. The main reason for this difference in calculation time is because the incremental demand between the pre-existing nodes

on the network that follows spatial progression is zero, eliminating the need for trip distribution calculation.

In this study, the airports were listed by G_i in descending order and added to the simulation to give benefits of preferential attachment to the airports with higher O-D demand. It should be noted that the airports were ordered by enplanements for the determination of theoretical baseline in Section 5.1 since enplanements include connecting passengers, which provides more information to the model and helps it to generate the emulated NAS better. With practical baseline, meaningful studies of real-life networks can be performed.

5.2.1 Before AF_i adjustment

A simulation was performed using spatial progression with the settings of the theoretical baseline other than the evolution path. The results are shown in Figure 93.



(a) Enplanements l-strip sum (b) Distance demand

Figure 93: Comparison of enplanements from simulation before AF_i adjustment and sDB1B of year 2000

5.2.2 After AF_i adjustment

AF_i of the selected airline hubs were increased as shown in Table 22. These airports were the hubs of major airlines in year 2000. The result of enplanements and distance

demand from the simulation compared the data from sDB1B of year 2000 is shown in Figure 94 and Figure 95. It should be noted that the AF_i settings are not unique and various combinations of AF_i can generate very similar results.

Table 22: AF_i of the selected hubs for practical baseline buildup

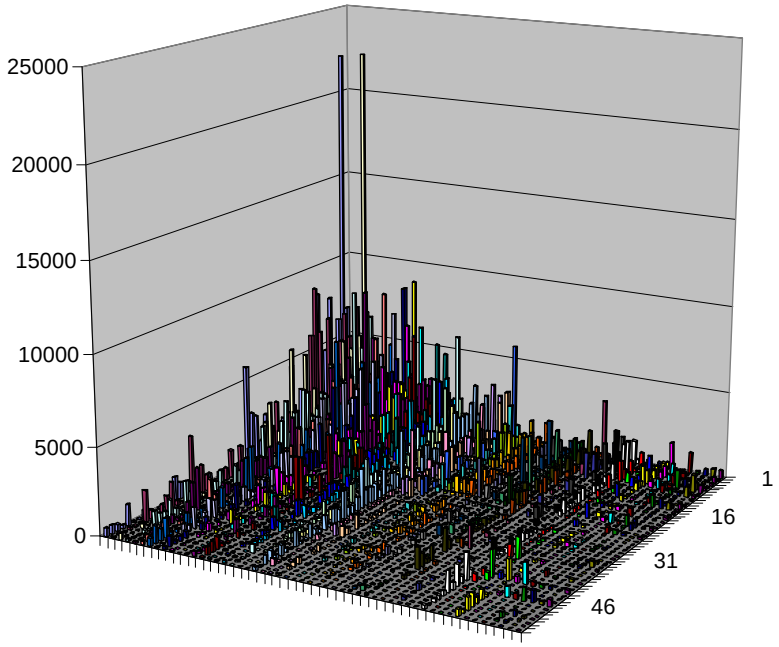
Airport	AF_i
ATL	0.14
CH_	0.04
CLT	0.2
CVG	0.29
DA_	0.13
DEN	0.09
DTW	0.06
HU_	0.24
MEM	0.39
MSP	0.11
PHX	0.11
PIT	0.12
SLC	0.18
STL	0.2

The comparison of enplanements from the practical baseline and sDB1B of year 2000 is given in Figure 96. It also supports the claim that the proposed framework is flexible enough to generate a network that is very similar to the NAS.

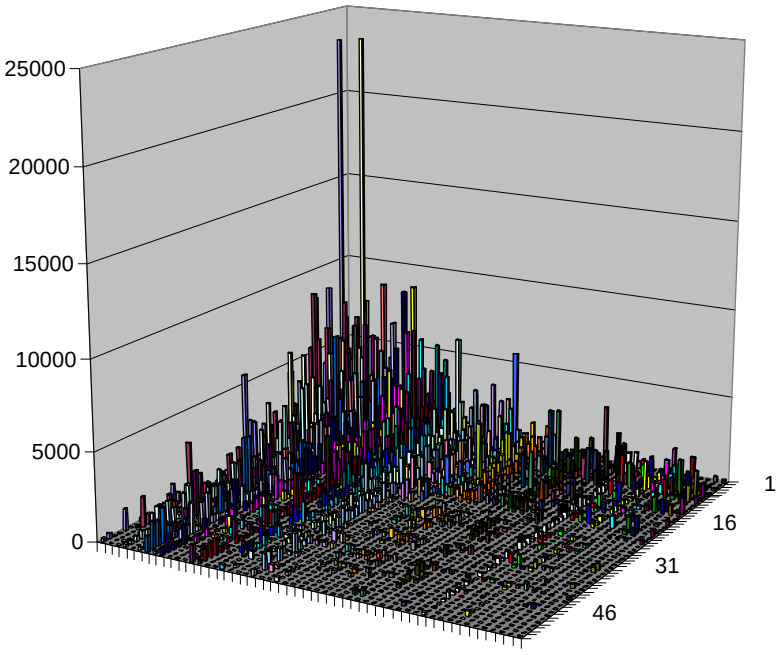
Figure 97 shows snapshots of the network on its evolution to its final stage. The last two subfigures (Figure 97e and Figure 97f) compares the enplanements from the simulation and from sDB1B. The visual comparison shows very close match. It should be noted that seven-color scale was used on the map, on which thicker and darker designates higher segmental enplanements.

Figure 98 shows comparison maps with segmental enplanements from the simulation and sDB1B. On these maps, seven-color scale was used based on the ranks of segmental enplanements.

Operations demand vs. distance by aircraft type was also analyzed. However,

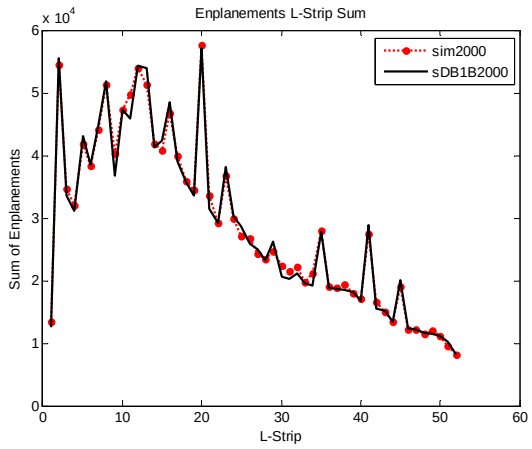


(a) Enplanements from simulation

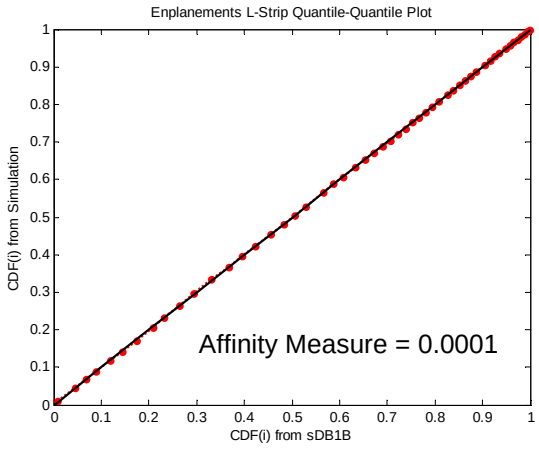


(b) Enplanements of year 2000 from sDB1B

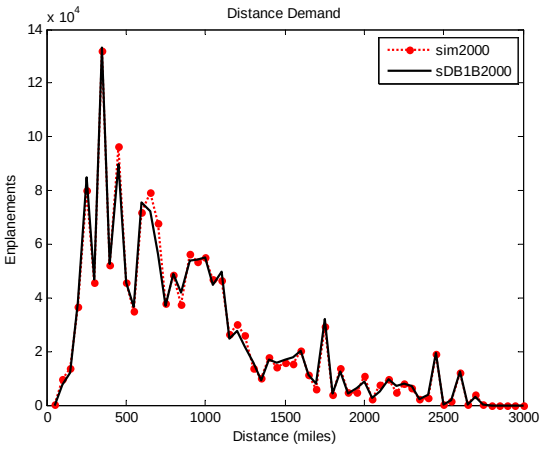
Figure 94: Comparison of enplanements from simulation with adjusted AF_i and sDB1B of year 2000



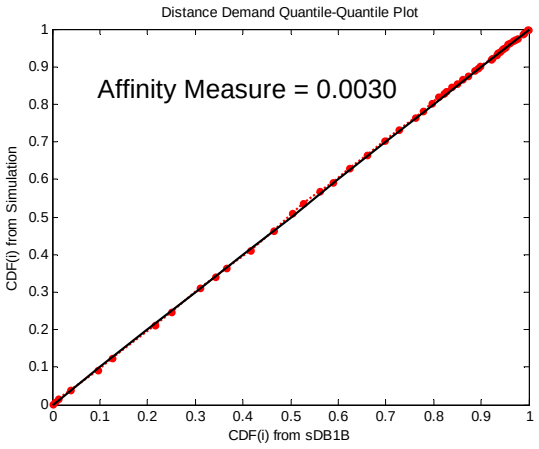
(a) Enplanements l-strip sum



(b) Enplanements l-strip QQ plot



(c) Distance demand



(d) Distance demand QQ plot

Figure 95: Comparison of enplanements l-strip sum and distance demand from simulation with adjusted AF_i and sDB1B of year 2000

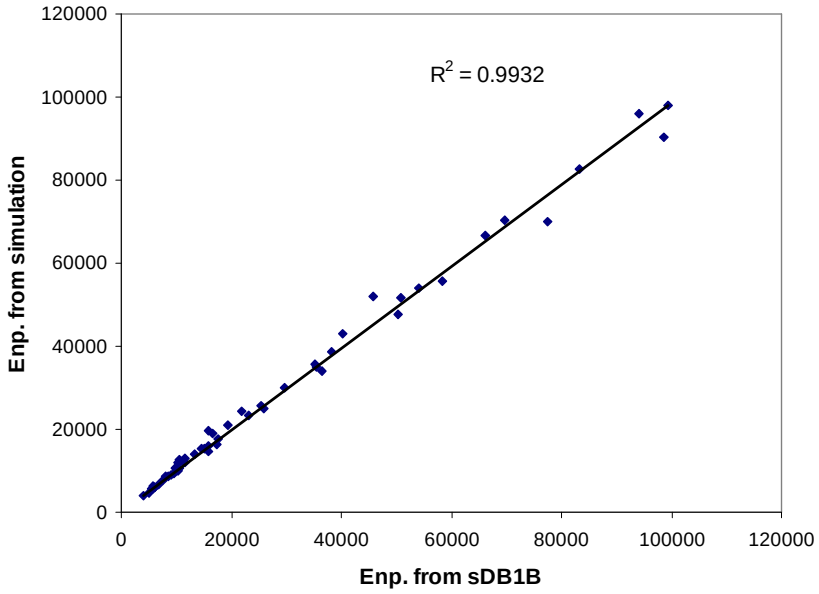


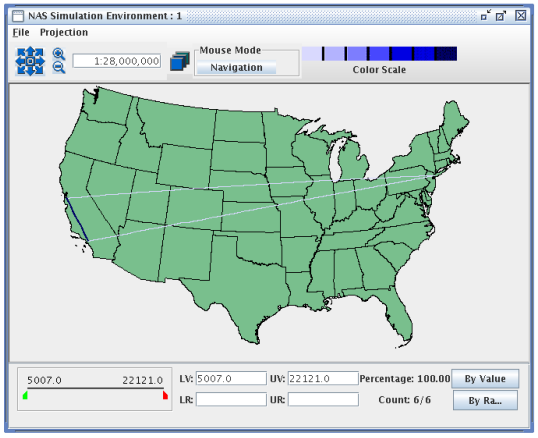
Figure 96: Prediction vs. actual: Enplanements from simulation with adjusted AF_i and sDB1B of year 2000

since capacity plays an important role in selecting aircraft types in the simulation, the five aircrafts were grouped into low- and high-capacity aircrafts, which are (TP, SRLC) and (SRHC, MR, and LR). The aggregated operations by aircraft type vs. the real-world data from T100D are compared in Figure 99 as distributions. The comparison of the low-capacity aircraft shows quite different characteristics. This discrepancy may have come from the following three reasons:

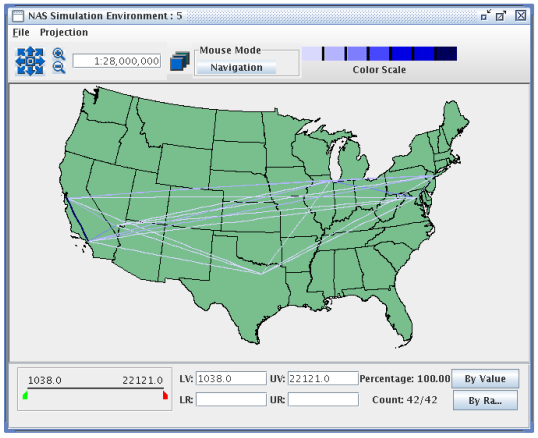
- Only one airline serves the whole O-D demand in the simulation while multiple airlines share the demand in the NAS. Therefore, the simulation tends to select high-capacity aircraft due to economies of scale.

- Segmentally optimal aircraft type is used in the simulation while sub-optimal aircraft types are used by airlines in the NAS to maximize the utilization of their resources.

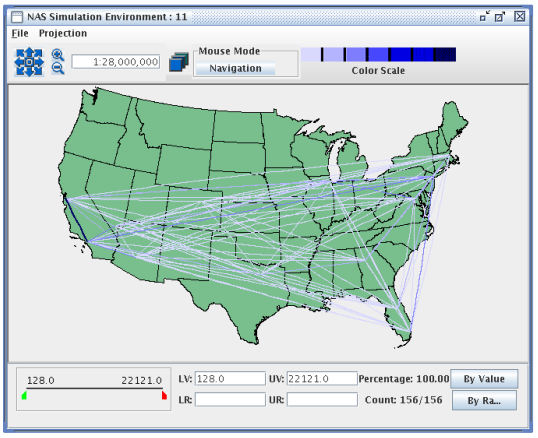
- Aircraft economical characteristics in this research were obtained by data gathering for variable cost and assumptions for fixed cost. Therefore, erroneous



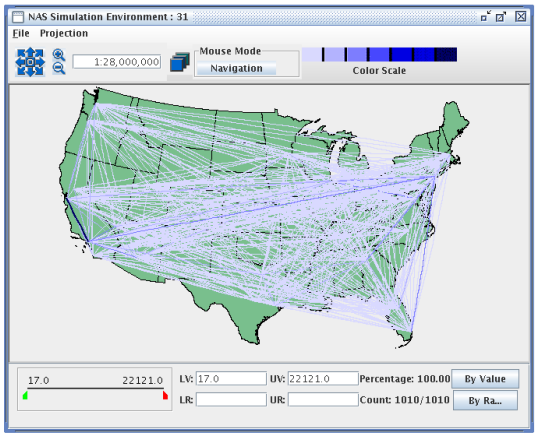
(a) Progression 1: 3 airports



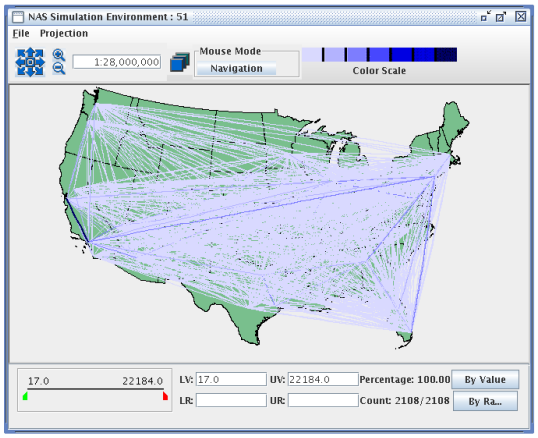
(b) Progression 5: 7 airports



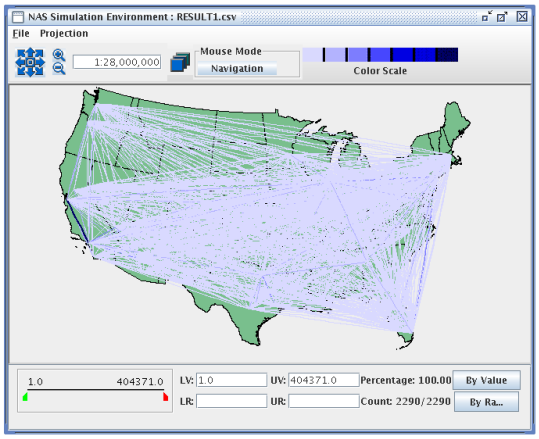
(c) Progression 11: 13 airports



(d) Progression 31: 33 airports

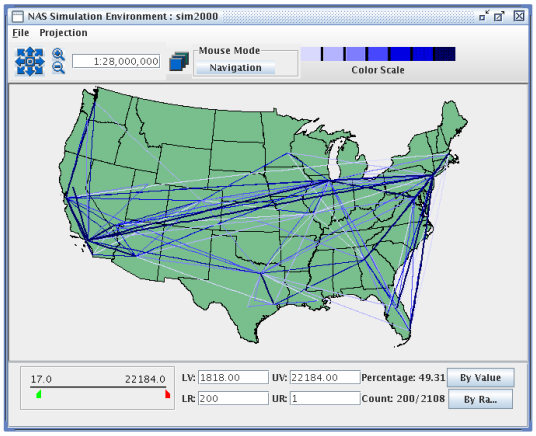


(e) Progression 51: 53 airports

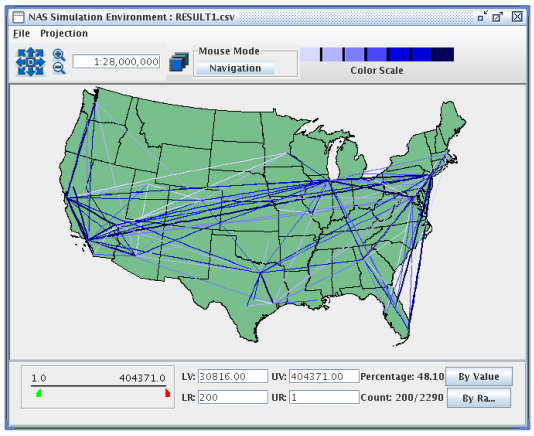


(f) Segmental enplanements from sDB1B: 53 airports

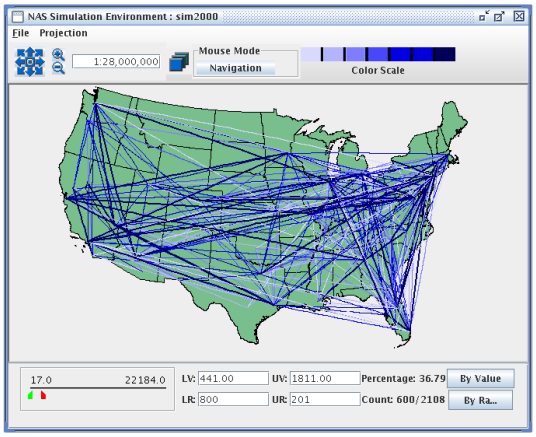
Figure 97: Snapshots of practical baseline progression



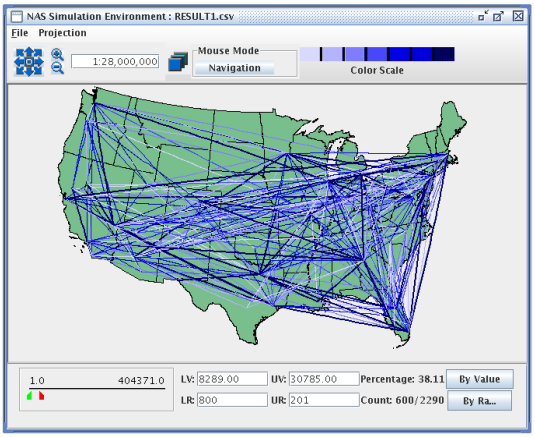
(a) Top 1^{st} to 200^{th} from simulation



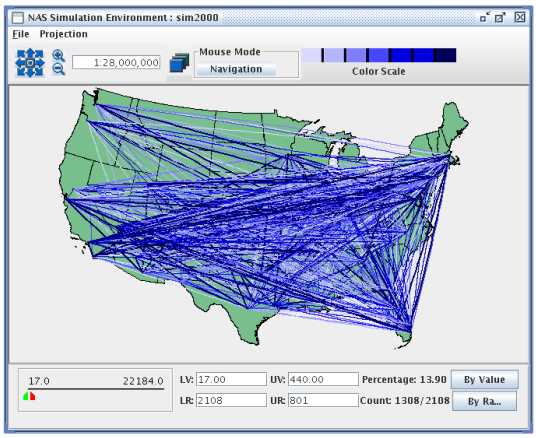
(b) Top 1^{st} to 200^{th} from sDB1B



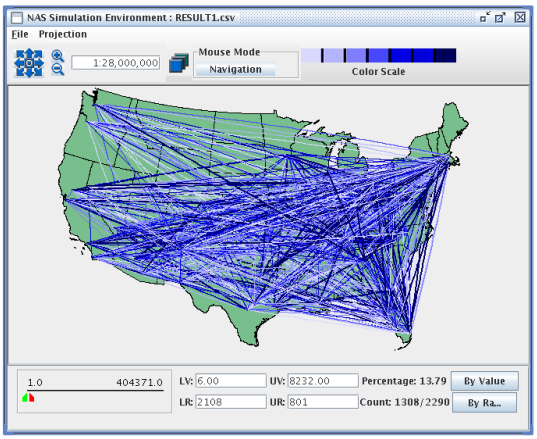
(c) Top 201^{st} to 800^{th} from simulation



(d) Top 201^{st} to 800^{th} from sDB1B

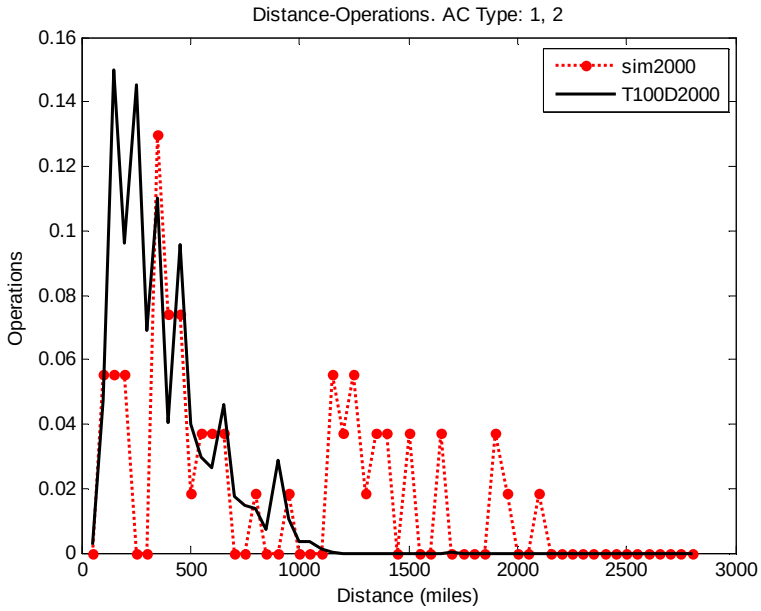


(e) Top 801^{st} to 2108^{th} from simulation

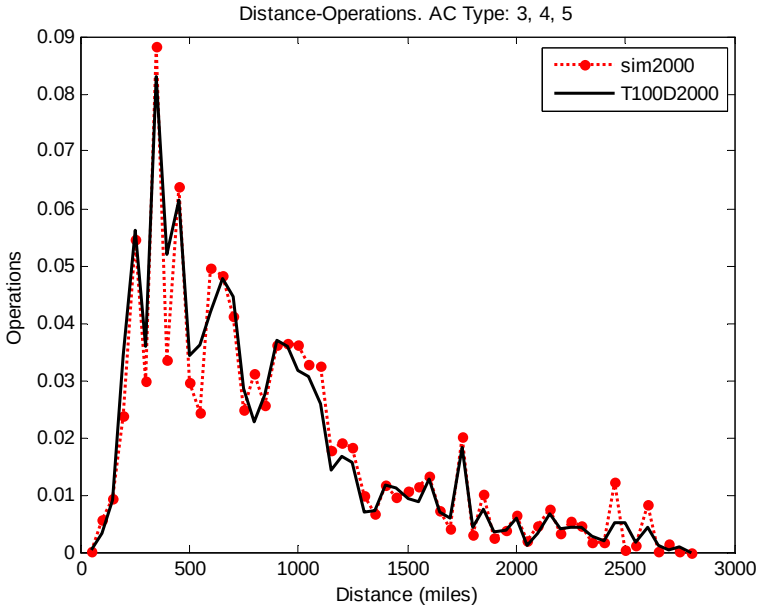


(f) Top 801^{st} to 2108^{th} from sDB1B

Figure 98: Segmental enplanements comparison by ranks



(a) Low-capacity (A/C Type 1,2)



(b) high-capacity (A/C Type 3,4,5)

Figure 99: Operations demand vs. distance by aircraft type

data entry and difference of accounting practices between airlines combined with inappropriate can alter the attractiveness of individual aircraft type.

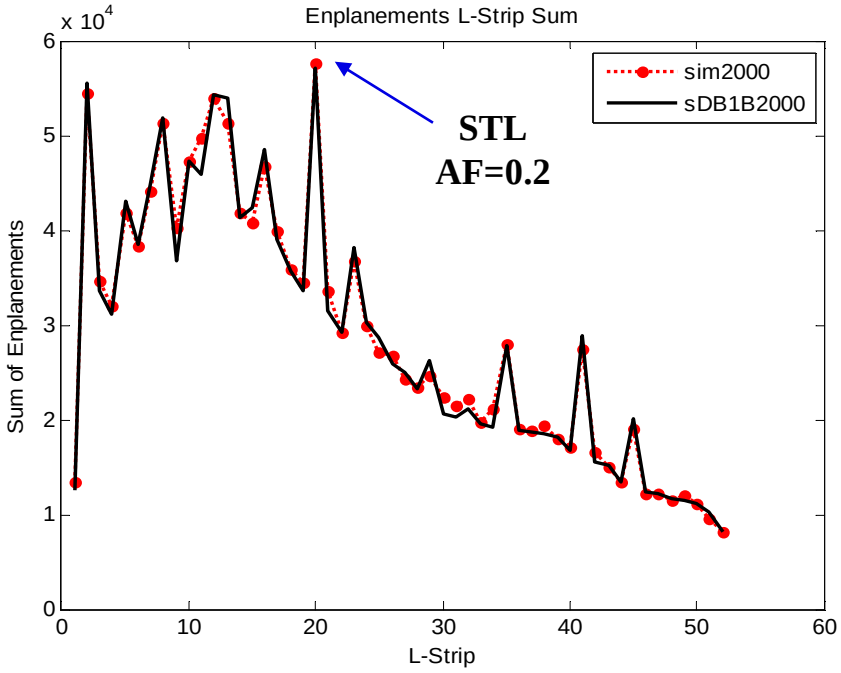
However, the comparison of the high-capacity aircraft shows that the practical baseline is very close to the NAS. It is because they service the majority of the demand in both simulation and the NAS.

5.3 Dehubbing of STL

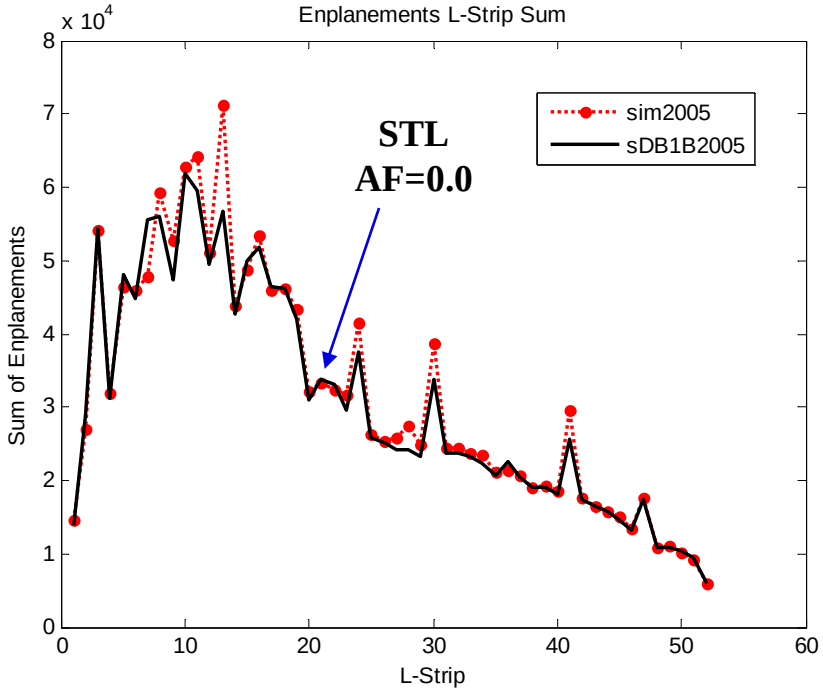
To demonstrate the prediction capability that is unique to the proposed framework, dehubbing scenario of Lambert-St. Louis International Airport (STL) was simulated. STL was the hub of Trans World Airlines (TWA) before it was acquired by American Airlines in 2001. TWA was one of the biggest domestic air carriers serving the most major cities. Therefore, STL has lost its role as a hub since then as it was witnessed in Figure 43c and Figure 44a in Chapter 4.

In this study, AF_i of the model for STL was decreased from 0.2 to 0.0, providing no incentives for hub connection at STL. The airports were ordered by G_i of year 2005 since the intent was to model th dehubbing of STL in year 2005. Figure 100a shows the practical baseline where $AF_{STL} = 0.2$ and Figure 100b shows the results from setting $AF_{STL} = 0$. The simulation results show that STL loses its hub role and other airports take extra load, which is similar to what happened in the NAS. However, it should be noted that the simulation cannot generate perfectly matching prediction since various strategic decisions are made in real air transportation network.

A visual comparison of the emulated and real NAS are shown in Figure 101 where thicker line describes higher enplanements. Figure 101a shows a route map of the NAS obtained from T100D for year 2000 when STL was used significantly by TWA and Figure 101b is the result of the simulation before setting attraction factor of STL neutral. Both maps depict very close markets in route-level. Figure 101c shows a route map of the NAS from T100D for year 2005 when STL lost its hub role completely.



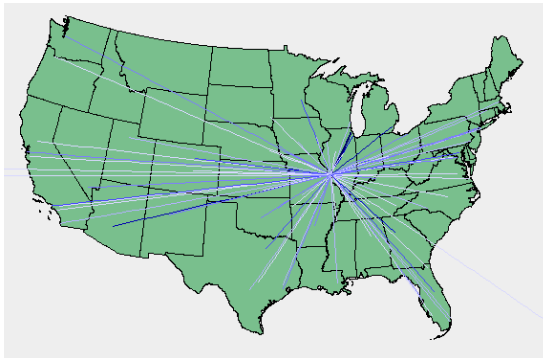
(a) Enplanements l-strip sum: before



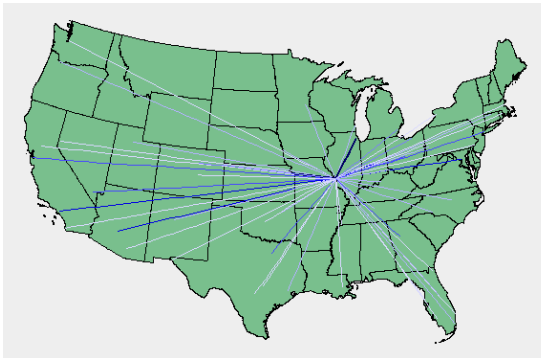
(b) Enplanements l-strip sum: after

Figure 100: Comparison of enplanements before/after setting $AF_i = 0$ for STL

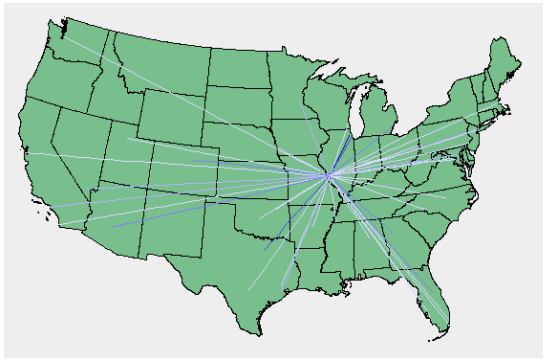
Figure 101d is the result of simulated dehubbing of STL. Visual comparison of the two maps shows that the result of the dehubbing scenario of STL closely matches what really happened on the NAS in route-level.



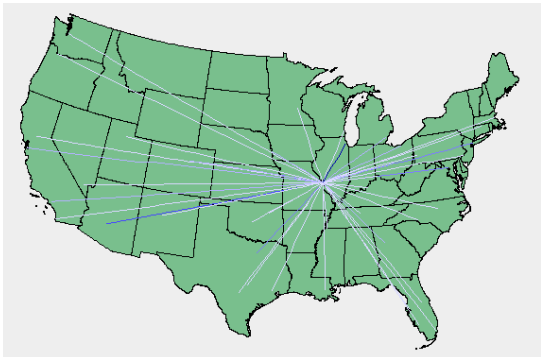
(a) Real: Route map from STL of year 2000 from T100D



(b) Simulation: Route map from STL before dehubbing



(c) Real: Route map from STL of year 2005 from T100D



(d) Simulation: Route map from STL after dehubbing

Figure 101: Segmental enplanements comparison by ranks

CHAPTER VI

CONCLUSIONS

6.1 Revisiting research questions and hypotheses

In this section, the research questions and hypotheses that were raised in the course of this research are summarized and how they were addressed is described. **Research Question 1** was raised after pointing out the problems of direct modeling of enplanements in H&S air transportation network study as exemplified in Figure 3, 4, 7, and 6 in Chapter 1. In an effort to differentiate intrinsic demand (τ_{ij}) from variable one (E_{ij}), the 'PACE' terminology and process was implemented. The decomposition of demand using PACE revealed the roles of airports on the air transportation network and produced (P_i) and attracted (A_i) demand demonstrated linear growth patterns. The linear pattern suggested that they are correlated with socioeconomic characteristics. Therefore, **Hypothesis 1** was made as below.

Research Question 1:
How can the intrinsic demand model be generated?
Hypothesis 1:
A demand prediction model that predicts intrinsic need can be built based on socioeconomic characteristics.

In the process to build models for P_i and A_i, many databases were overviewed and DB1B was selected as the best candidate from which the intrinsic demand characteristics can be extracted. However, due to the built-in error of DB1B, it was processed to collect only symmetric 2- and 4-coupon data with single-coupon trips paired-up according to the true O-D demand ratios (i.e., t_{ij} to t_{ji}) obtained from the symmetric 2- and 4-coupon data. The final product was named sDB1B and it was validated

against T100D to verify they are acceptably similar. With sDB1B at hand, the demand models to predict P_i and A_i were pursued. As an approach, a 'Push-and-pull' model that resembles the gravity law with county-level influences residing at their population centroid was tested. Even with the lack of available county-level socioeconomic data, this simple model generated very promising results, which was enough to prove the **Hypothesis 1**.

The importance of modeling intrinsic demand was demonstrated using the two-factor FRATAR model introduced in Section 4.1.6, which scaled up the baseline τ_{ij} matrix with P_i and A_i as growth factors. Even though the NAS went through severe structure changes as observed in the investigation of the NAS in Section 3.1, it was possible to generate highly accurate τ_{ij} matrix of year 2005 from the one of year 1997.

After the intrinsic demand modeling process was set up, the next question was about how to generate airline network model. The literature search in the field of airline network modeling revealed that current practices treat the network as equilibrium system, not evolutionary one. Therefore, **Research Question 2** about the evolutionary nature of the airline network was raised. Since there was enough information about the evolutionary nature of the NAS, an effort was made to preserve the structure of the evolving NAS when building an airline network model. Since majority of complex networks can be explained with preferential attachment as overviewed in Section 2.2, **Hypothesis 2** was made as below.

Research Question 2:

Is airline network evolutionary?

Hypothesis 2:

Airline network is evolutionary and can be explained with preferential attachment.

Now the next question (**Research Question 2-1**) was about how to implement preferential attachment in an airline network model. Four hypotheses were made as

the study progressed with the ideas from the literature search (**Hypothesis 2-1-1**), database study (**Hypothesis 2-1-2**), and pure reasoning (**Hypothesis 2-1-3** and **Hypothesis 2-1-4**), which are summarized below.

Research Question 2-1:

How can preferential attachment be implemented in airline network?

Hypothesis 2-1-1:

Airline network can be evolved by adding highly demanded vertices first and evolving them to a certain maturity before adding the rest within a very short timeline.

Hypothesis 2-1-2:

Many airports in the NAS deviates from general power-law. By introducing attraction factors and fine-tuning the attractiveness on a growing network, the real transportation network can be emulated.

Hypothesis 2-1-3:

Probabilistic choice model can capture the non-optimal route choice behavior on the NAS.

Hypothesis 2-1-4:

The growth of NAS can be emulated by spatial and chronological evolution scheme.

Based on these four hypotheses, an evolutionary airline network model was composed as described in Section 3.3 and Section 4.2. The proposed evolutionary air transportation network model tries to capture non-optimal behavior of the network by utilizing multinomial logit model in its core. The number of airports and the O-D demand at each iteration is determined by the predetermined evolution path. The results from the theoretical baseline study in Section 5.1 showed that the evolution path that resulted in a network similar to sDB1B has convex shape. The convex path adds small number of airports into the simulation and evolve them to form the initial structure of the core network and inserts the rest of the airports in a very short

time period. Therefore, the study showed that **Hypothesis 2-1-1** was appropriate. Also, a point was made that overall evolution path for the whole NAS might have 'S' shape curve as described in Figure 92. It was also shown that evolution path works as a mechanism that compensates incompleteness of airline network model. The study of practical baseline determination in Section 5.2 showed that the NAS can emulated very closely by fine-tuning the model with attraction factors. This proves **Hypothesis 2-1-2** stating that the deviation from general power-law of the airports in the NAS can be modeled with the attraction factors. Since the **Hypothesis 2-1-1** and **Hypothesis 2-1-2** were proved, the rest two hypotheses (**Hypothesis 2-1-3** and **Hypothesis 2-1-4**) that were used to generate the simulation framework are indirectly proved to be appropriate.

6.2 Contribution

The inherent complexity of the NAS made it difficult for the researchers to understand the dynamics existing in the system. In this research, intrinsic demand was separated from the network enplanements and active design algorithm was created for transport architect, who wants to investigate the impact of network topology changes. The research was performed with a new perspective for both demand and network modeling. Since the current state-of-the-art airline network models do not include evolutionary network modeling scheme and topological usage changes of the airline network are not treated explicitly, this new active design algorithm can be a good tool for a transport architect who desires to expand his/her design space to include topological changes of the airline network. The main contribution of this research is in developing a method for demand- and network-centric analysis and viewing the airline network from evolutionary point of view. The contributions are summarized below.

1. Decoupled variant and invariant element from enplanements demand (Section 4.1)

- Demonstrated that demand modeling should focus on invariant element
- Future true O-D demand matrix was successfully generated when there was significant disturbances in the topological structure of the NAS

2. Created evolution-based airline network model (Section 4.2)

   - New modeling approach that preserves the evolutionary nature of the NAS was created
   - Developed a design process for evolutionary air transportation networks where the process starts with intrinsic demand (t) matrix, the growth of it is performed using the predicted P_i and A_i utilizing two-factor FRATAR, and the variable demand is modeled on a evolutionary network model.
   - An architect of airline network can perform network studies by changing the attraction factors for hub connections, extending conventional study with fixed topology to topologically unconstrained ones.

6.3 Future work

When sDB1B was constructed, only symmetric trips are gathered. However, if intelligent data gathering techniques can be used to extract more information such as multiple destination trips from DB1B, more fruitful results are expected to be obtained.

It was shown that two-factor FRATAR process could be used to scale baseline t demand matrix. When more accurate future P_i and A_i can be obtained by building a model with better socio-economical data, it is expected that accurate future t matrix can be obtained utilizaing the two-factor FRATAR model. Also, in the development process of P_i and A_i model, only single point in time (year 1997) was used. Enhanced models can be built if multiple calibration points are available. Other forms of the models may also be tried to improve the accuracy, such as in analogy of source-sink

model in fluid mechanics or magnetic field model in physics.

Even though the evolutionary air transportation network model demonstrated prediction capability using the dehubbing of STL (See Section 5.3, the functional relationship between the attraction factors (AF_i) and the characteristics of airports in real life was not determined. Also, AF_i sets used to generate the practical baseline in Section 5.2 is not unique. Various combinations of AF_i adjustment can generate similar results. What and how to select the best combination still remains as a question. A research in this area will bring a great deal of benefit to the prediction of future air transportation networks.

The methods and results obtained from this research may also be extended to analyze itinerary-centric aspect of the airline network. The extension of this research into itinerary-centric analysis is important since it gives focus to passengers and passenger delays are more significant than flight delays. Itinerary-centric analysis will open the door for addressing the problem of itinerary-centric delay propagation, which may include the whole passenger itinerary of complete curb-to-curb travel. The analysis of the curb-to-curb delay propagation in the NAS will provide valuable input for the enhancement of the NTS development. Also this research can be extended to link aircraft design problems with strategic airline network development problems to come up with the best design of the aircraft for the anticipated airline network by concurrently optimizing vehicle design and network flow as suggested by Taylor and de Weck (2007) and Mane et al. (2007).

***

APPENDIX A

GENERAL TRANSPORTATION DEMAND FORECASTING

In this section, the general transportation demand forecasting processes are overviewed. General transportation demand forecasting involves estimating the number of travelers and vehicles in single or multi mode cases. Examples of forecasts include the number of cars on a specific routes on a highway, number of airplanes landing per hour at an airport, or the number of passengers at a seaport. Usually the process starts with building a demand model by finding the relationship between current traffic data and socio-economic characteristics of the considered regions, for example population, income, accommodation and food sales, employment rate, travel cost, etc. Inserting futures of those socio-economic values in the model results in the prediction. Transportation demand forecasts are important in developing capacity and growth plans, to calculate economic and environmental impact studies, etc.

In this section, the characteristics of selected demand estimation models, trip distribution models, and choice models are introduced.

A.1 Traditional four-step process

These steps usually consists of four steps starting from the gross demand estimation and ending with determining routes for each travel as below.

1. Macroscopic demand estimation: Travel demand in each zone is estimated from the macroscopic demand estimation models.

2. Trip distribution: Trip demand is distributed among zones in the network using the demand distribution model. The results become the gross demand on each

origin-destination (OD) pair. The most popular models are gravity model and entropy model.

3. Transportation mode choice: Transportation modes (auto, bus, train,or air) are selected for each OD pair. The most popular discrete choice model for this purpose is of the logit form developed by Daniel McFadden.

4. Travel route assignment: The specific route of the selected mode of transportation from origin to destination is selected.

A.2 Demand estimation models

Airline operations research often involves prediction of the air travel demand. Macroscopic demand models are used when the aggregate or regional development levels of transportation service need to be described. The number of travelers and the number of takeoffs at a region are typical usages of the macroscopic models. On the other hand, microscopic models are used to assess detailed properties of transportation services. Passenger traffic along a specific route and the number of passengers in each mode of transportation are typical usages of microscopic models.

A.2.1 Gross travel demand estimation model

There exist many macroscopic demand models from simple linear models to complicated socio-economic models. The two most common ways of differentiating the demand models are time based models and socio-economy based demand models.

The time based models predict demand volume using previous trends. Since it is generally agreed that there exist relationship between the transportation system demand and the socio-economic characteristics of a considered market, time based models don't capture the cause and effect relationship. Some of the most widely used time based demand models are given below. In these models, all the variables except time need to be estimated from the existing demand data.

$y = kt + m$ (Linear)

$y = k + a \times \ b^t$ (Exponential)

$y = k \times a^{b^t}$ (Gompertz Curve)

$y = \frac{k}{1+be^{-at}}$ (Pearl-Reed curve, Logistic curve)

Another way of constructing a demand model is to connect certain socio-economic characteristics and the demand level of a certain transportation mode. Socio-economy based models use historical data on the number of people, income, number of tourists, and other suitable socio-economic characteristics to capture the relationship behind the actual demand. Also, the speed, cost, and other characteristics of the transportation services can be included in the model. Historical data collection and traveler surveys are the most common way of obtaining the required information. With the data, the model is calibrated and the transportation system demand is forecasted by changing values of the socio-economic variables to the expected level. The general form of the socio-economy based demand model is below.

$$Y_t = a \prod_{i=1}^{m} S_{it}^{b_i} \prod_{j=1}^{n} T_{jt}^{c_j}$$

Where,

m : total number of socio-economic characteristics.

n : total number of transportation characteristics.

Y_t : number of air passengers in year t.

S_{it} : i^{th} socio-economic variable value in year t.

T_{jt} : j^{th} transportation characteristic value in year t.

a, b_i, c_j : parameters.

A.2.2 Intercity travel demand estimation model

Intercity travel demand estimation models can be categorized into two types depending on the distance of the travel — short and medium haul and long haul travel. In the short and medium haul travel cases, travelers are sensitive to the price and the level of convenience of other modes of transportation to their specific needs. Therefore, multimode models need to be used that include many competitive modes of transportation, e.g. personal auto, bus, and trains. For long haul travel, airline is the dominant choice of transportation and other services rapidly lose attractiveness due to too much travel time as the distance increases.

To assess new transportation service effectiveness, a demand model should have an ability to predict percentage share of each modes of transportation services in the regions considered. The abstract mode model developed by Quandt and Baumol (1966) can be used for the purpose and it can estimate the number of passengers that would choose a new mode of transportation (Teodorović, 1988). The abstract model assumes that a traveler's transportation mode choice depends on both relative and best characteristics of the systems. The abstract model in the original form is shown below.

$$T_{kij} = \alpha_0 P_i^{\alpha_1} \cdot P_j^{\alpha_2} \cdot Y_i^{\alpha_3} \cdot Y_j^{\alpha_4} \cdot M_i^{\alpha_5} \cdot M_j^{\alpha_6} \cdot N_{ij}^{\alpha_7} \cdot (H_{ij}^b)^{\beta_0} \cdot (H_{kij}^r)^{\beta_1} \cdot (C_{ij}^b)^{\gamma_0} \cdot (C_{kij}^r)^{\gamma_1} \cdot (D_{ij}^b)^{\delta_0} \cdot (D_{kij}^r)^{\delta_1}$$

Where,

T_{kij} : the number of trips between city i and city j by transportation mode k.

P_i, P_j : the populations in cities i and j

Y_i, Y_j : the average national income in cities i and j

M_i, M_j : the percentage of the population employed in industry in cities i and j

N_{ij} : the number of different transportation modes operating between cities i and j

H_{ij}^b : the least possible travel time between cities i and j

H^r_{kij} : relative travel time between cities i and j by transportation mode k

C^b_{ij} : the least possible travel costs between i and j

C^r_{kij} : relative travel costs between i and j by transportation mode k

D^b_{ij} : the greatest frequency (number of departures) between cities i and j

D^r_{kij} : relative frequency between cities i and j by transportation mode k

$\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \beta_0, \beta_1, \gamma_0, \gamma_1, \delta_0, \delta_1$: parameters estimated statistically

The equation can be calibrated to data by applying logarithms on both sides and using multiple regression methods with the linearized form. By putting the characteristics of the new transportation system into the calibrated model, the number of passengers who will use the new transportation system can be estimated.

A.3 Trip distribution models

A.3.1 Entropy model

A.3.1.1 Entropy theory in transportation systems research

Information theory was started by Claude E. Shannon with his paper "A Mathematical Theory of Communication" (Shannon, 1948). In the paper, Shannon devised a measure of uncertainty and provided a mathematical background of entropy theory in communication networks. The information-theoretic entropy is measured to solve a communication problem of efficiently transferring information through noisy communication channels. His idea evolved into two main branches of research — Information theory and coding theory.

The information theoretic approach can be applied to a trip distribution problem on a transportation network. Trip distribution models generate trips between individual cities with the travel demand established in the previous steps. Even though there are many other trip distribution models frequently used in airline operations research, entropy models are chosen for review due to its solid mathematical explanation.

Using the notation of Teodorović (1988) and denoting the number of trips from

origin i to destination j to be T_{ij}, flight frequencies of an air transportation system can be represented as a square matrix. Figure 102 illustrates this process. The column on the right side of the travel matrix is the assessed outgoing trip demand from each city. The row at the bottom is the assessed incoming trip to each city. With the demand information, the trip from city i to city j (T_{ij}) should be obtained using the trip distribution model.

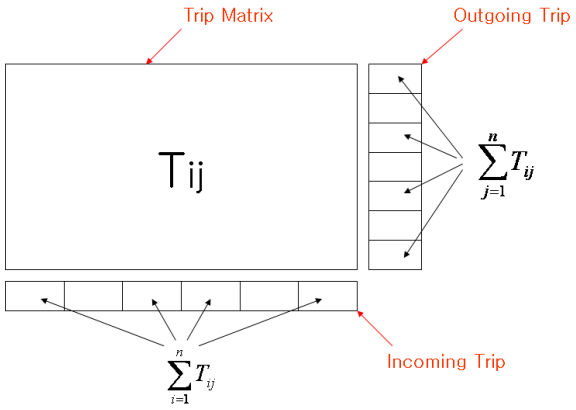


Figure 102: Trip distribution process

To set up the equations for a trip distribution model, the following equalities are introduced by defining T_{ij}: the number of trips from origin i to destination j, a_i: the number of trips generated by node i, b_j: the number of trips attracted to node j, and k: the total number of trips.

$$a_i = \sum_{j=1}^{n} T_{ij} \qquad b_j = \sum_{i=1}^{m} T_{ij} \qquad k = \sum_{i=1}^{m} a_i = \sum_{j=1}^{m} b_j$$

$$u_i = \frac{a_i}{k} \qquad v_j = \frac{b_j}{k} \qquad p_{ij} = \frac{T_{ij}}{k}$$

$$\sum_{i=1}^{m} u_i = \sum_{j=1}^{m} v_j = \sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} = 1$$

With the notations given above, the entropy of the transportation network can be

defined as:

$$H = \sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} ln(p_{ij})$$

The probabilities of trips being made between all cities in the network can be calculated by finding the condition for maximum entropy.

A.3.1.2 Proportional model

The proportional model is formulated by enforcing the basic entropy model to match the number of outgoing and incoming trips on the network. The equations that need to be solved to calculate the probabilities p_{ij} are given below.

$$H = \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} ln(p_{ij})$$

$$\sum_{j=1}^{m} p_{ij} = u_i \quad \forall i$$

$$\sum_{i=1}^{m} p_{ij} = v_j \quad \forall j$$

$$\sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} = 1$$

To solve the equations, the Lagrangian function F and Lagrange multipliers α, β, and γ are introduced to find the solution.

$$F = -\sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} ln p_{ij} - \sum_{i=1}^{m} \alpha_i (\sum_{j=1}^{m} p_{ij} - u_i) - \sum_{j=1}^{m} \beta_j (\sum_{i=1}^{m} p_{ij} - v_j) - \gamma (\sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} - 1) \tag{13}$$

To generate the equations that can be solved for p_{ij}, partial derivatives are taken on the above equation and each of them are set equal to zero. For the solution, the following equations need to be solved together.

$$\frac{\partial F}{\partial p_{ij}} = -lnp_{ij} - 1 - \alpha_i - \beta_j - \gamma = 0 \tag{14}$$

$$\frac{\partial F}{\partial \alpha_i} = -\sum_{j=1}^{m} p_{ij} + u_i = 0 \tag{15}$$

$$\frac{\partial F}{\partial \beta_i} = -\sum_{i=1}^{m} p_{ij} + v_j = 0 \tag{16}$$

$$\frac{\partial F}{\partial \gamma} = -\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij} + 1 = 0 \tag{17}$$

From (14) , p_{ij} is represented as below

$$p_{ij} = e^{-(1+\alpha_i+\beta_j+\gamma)} \tag{18}$$

Substituting (18) with p_{ij} in (15), (16), and (17) result,

$$u_i = e^{-(1+\alpha_i+\gamma)} \sum_{j=1}^{m} e^{-\beta_j} \tag{19}$$

$$v_j = e^{-(1+\beta_j+\gamma)} \sum_{i=1}^{m} e^{-\alpha_i} \tag{20}$$

$$e^{-(1+\gamma)} \sum_{i=1}^{m}\sum_{j=1}^{m} e^{-\alpha_i} e^{-\beta_j} = 1 \tag{21}$$

Now, multiplying (19) and (20), and simplifying the result with (21) gives the equation for p_{ij}.

$$u_i \cdot v_j = e^{-2(1+\gamma)} e^{-\alpha_i} e^{-\beta_j} \sum_{i=1}^{m}\sum_{j=1}^{m} e^{-\alpha_i} e^{-\beta_j} = e^{-(1+\alpha_i+\beta_j+\gamma)} = p_{ij} \tag{22}$$

This equation estimates that the probability of travel between the two cities are equal to the product of the probability of travel generation of the city i and the

probability of travel attraction of the city j. From the definition, the frequency of the travel between the city i and the city j is estimated as below.

$$T_{ij} = \frac{a_i \cdot b_j}{k} \tag{23}$$

A.3.1.3 Mean trip length model

The mean trip length model is formulated when the cost factor is included in the proportional model. The name originated from the fact that the trip length is often included in the cost calculation. By denoting c_{ij} as the trip cost from city i to city j, the total cost of a transportation network, TC, can be represented as below,

$$TC = \sum_{i=1}^{m} \sum_{j=1}^{m} T_{ij} c_{ij}$$

Since $p_{ij} = T_{ij}/k$, the average cost per passenger of a transportation network, $AC = TC/k$, can be represented as below,

$$AC = \sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} c_{ij}$$

By incorporating the cost information into the proportional model, equations for the mean trip length model is formulated. The resulting trip probabilities between city i and city j, p_{ij}, are obtained by finding the maximum entropy (H) that satisfies the constraints. The entropy equation and constraints for the situation are given below.

$$H = \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} ln(p_{ij})$$

$$\sum_{j=1}^{m} p_{ij} = u_i$$

$$\sum_{i=1}^{m} p_{ij} = v_j$$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij} = 1$$

$$\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}c_{ij} = AC$$

To solve the equations, Lagrangian function F and Lagrange multiplier α, β, γ, and ψ are introduced.

$$\begin{aligned} F &= -\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij} ln p_{ij} - \sum_{i=1}^{m} \alpha_i \left(\sum_{j=1}^{m} p_{ij} - u_i \right) - \sum_{j=1}^{m} \beta_j (\sum_{i=1}^{m} p_{ij} - v_j) \\ &- \gamma(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij} - 1) - \psi(\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}c_{ij} - AC) \end{aligned}$$

To generate the equations that can be solve for p_{ij}, partial derivatives are taken on the above equation and each of them are set equal to zero. For the solution, the following equations need to be solved together.

$$\frac{\partial F}{\partial p_{ij}} = -ln p_{ij} - 1 - \alpha_i - \beta_j - \gamma - \psi c_{ij} = 0 \tag{24}$$

$$\frac{\partial F}{\partial \alpha_i} = -\sum_{j=1}^{m} p_{ij} + u_i = 0 \tag{25}$$

$$\frac{\partial F}{\partial \beta_i} = -\sum_{i=1}^{m} p_{ij} + v_j = 0 \tag{26}$$

$$\frac{\partial F}{\partial \gamma} = -\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij} + 1 = 0 \tag{27}$$

$$\frac{\partial F}{\partial \psi} = -\sum_{i=1}^{m}\sum_{j=1}^{m} p_{ij}c_{ij} + AC = 0 \tag{28}$$

From (24), p_{ij} is represented as below.

$$p_{ij} = e^{-(1+\alpha_i+\beta_j+\gamma+\psi c_{ij})} \tag{29}$$

Substituting (29) with p_{ij} in (25), (26), (27), and (28) result,

$$u_i = e^{-(1+\alpha_i+\gamma)} \sum_{j=1}^{m} e^{-\beta_j - \psi c_{ij}} \tag{30}$$

$$v_j = e^{-(1+\beta_j+\gamma)} \sum_{i=1}^{m} e^{-\alpha_i - \psi c_{ij}} \tag{31}$$

$$e^{-(1+\gamma)} \sum_{i=1}^{m} \sum_{j=1}^{m} e^{-\alpha_i - \beta_j - \psi c_{ij}} = 1 \tag{32}$$

$$AC = e^{-(1+\gamma)} \sum_{i=1}^{m} \sum_{j=1}^{m} c_{ij} e^{-\alpha_i - \beta_j - \psi c_{ij}} \tag{33}$$

An iterative method can be used to solve the equations above for p_{ij}. To simplify the equations, substitutions are made as below.

$$P_i = e^{-1-\alpha_i-\gamma}$$

$$Q_{ij} = e^{-\psi c_{ij}}$$

$$R_j = e^{-\beta_j}$$

After the substitutions, p_{ij} is represented as $p_{ij} = P_i Q_{ij} R_j$.

The equations to be solved become the following.

$$u_i = P_i \sum_{j=1}^{m} Q_{ij} R_j \tag{34}$$

$$v_j = R_j \sum_{i=1}^{m} P_i Q_{ij} \tag{35}$$

$$AC = \sum_{i=1}^{m} \sum_{j=1}^{m} c_{ij} P_i Q_{ij} R_j \tag{36}$$

In the first iteration, the values for ψ and β_j are assumed giving the starting values for Q_{ij} and R_j. The value of Q_{ij} doesn't change until the converged solution is obtained. The recommended value for ψ is $1/AC$ and AC is arbitrarily selected initially (Potts and Oliver, 1972).

The value of the first R_j is commonly set to be, $R_j^{(1)} = v_j$. Then the $P_i^{(1)}$ in the first iteration becomes as below.

$$P_i^{(1)} = \frac{u_i}{\sum_{j=1}^{m} Q_{ij} R_j^{(1)}} \tag{37}$$

The above result can be used to calculate $R_j^{(2)}$ in the second iteration as below.

$$R_j^{(2)} = \frac{v_j}{\sum_{i=1}^{m} P_i^{(1)} Q_{ij}} \tag{38}$$

The iteration continues until the convergence criterion is satisfied. The equations on the k^{th} iteration are as below.

$$R_j^{(k)} = \frac{v_j}{\sum_{i=1}^{m} P_i^{(k-1)} Q_{ij}}$$

$$P_i^{(k)} = \frac{u_i}{\sum_{j=1}^{m} Q_{ij} R_j^{(k)}}$$

After each iteration, p_{ij} is calculated using the equation below.

$$p_{ij}^{(k)} = P_i^{(k)} Q_{ij} R_j^{(k)}$$

Iteration continues until the following stopping condition is met.

$$\max_{ij} \{ |p_{ij}^{(k)} - p_{ij}^{(k-1)}| \} \leq \epsilon_1$$

When iteration ends, a confirmation test should be performed to check the rest of the conditions are met using the equation below.

$$AC = \sum_{i=1}^{m} \sum_{j=1}^{m} c_{ij} P_i Q_{ij} R_j$$

$$\sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} = 1$$

The value of AC is changed and the process is repeated until the assumed AC and the calculated AC after iteration matches.

$$|AC - AC'| \leq \epsilon_2$$

A.3.2 Gravity model

Gravity model is the most widely used and documented trip distribution in transportation research community (Garber and Hoel, 2002). As in the law of gravity, the travel demand between origin and destination is directly proportional to the produced demand quantity of the origin and the attracted demand quantity of the destination. Mathematical form of the gravity model is expressed in Equation 39, reproduced from Garber and Hoel (2002).

$$T_{ij} = P_i \left[\frac{A_j \cdot F_{ij} \cdot K_{ij}}{\sum_j A_j \cdot F_{ij} \cdot K_{ij}} \right] \tag{39}$$

where

T_{ij} = number of trips from zone i to zone j

P_i = produced trips in zone i

A_j = attracted trips to zone j

F_{ij} = an inverse function of travel time i

K_{ij} = socioeconomic adjustment factor for interchange ij

A.3.3 Growth factor model: FRATAR Model

Growth factor models can be used when there are data available for O-D matrix for base year and the trip generation values for each zone for the future year. Because growth factor models scales up the base O-D matrix, they cannot be used to forecast traffic between zones where there is no base demand. Also, they cannot reflect travel friction changes such as travel time between zones. FRATAR method is the most popular growth factor model, which calculates future trip estimation by proportioning the relative growth between zones. Equation 40 shows the the mathematical form of FRATAR model, reproduced from Garber and Hoel (2002).

$$T_{ij} = (t_i \cdot G_i) \left[\frac{t_{ij} \cdot G_j}{\sum_x t_{ix} \cdot G_x} \right] \tag{40}$$

where

T_{ij} = number of trips estimated from zone i to zone j

t_i = present trip generation in zone i

G_x = growth factor of zone x

T_i = $t_i G_i$ = future trip generation in zone i

t_{ix} = number of trips between zone i and other zones x

t_{ij} = present trips between zone i and zone j

G_j = growth factor of zone j

A.4 Choice models: Logit model

In transportation research, there are many choice models that can be used to model an individual traveler's decision on selecting a certain mode of transportation. Choice models almost always deal with personal preferences like comforts, speed, cost, etc. Even though these individual preferences obviously affect travelers' behavior, it is difficult to assess them. The most widely accepted method for this problem is to poll some of the travelers and gather information on them to build a choice model. When polling non-demographic factors, typically a scale is used to capture the traveler's

opinion. Travelers are asked to place an 'X' mark on the scale and the numerical value of each traveler's response is obtained by measuring the relative position of the 'X' mark on the scale. An example of the scale for safety evaluation is shown in Figure 103.

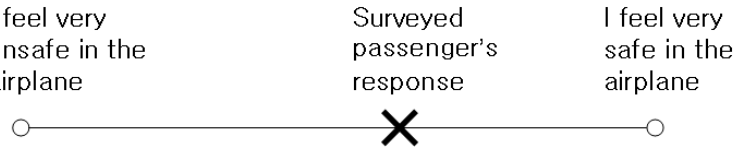


Figure 103: Example: Safety evaluation scale. [Source: Teodorović (1988)]

There are mainly two different types of choice models — Deterministic models and stochastic models. Deterministic models produce the same choice for each individual if the individual is in the same situation as others. Stochastic models add stochastic properties that allow a traveler to make different choices even when the captured situation is the same. It is known that stochastic models better model real life situations because there are many random effects that cannot be modeled into the choice model. Also the fact that travelers have different evaluation criteria for the utility of a specific transportation makes stochastic models a better choice. Another reason can be that travelers don't have the same level of information as others in many cases or they just don't choose the optimal solution even when they clearly see it.

Stochastic choice models consist of a deterministic part and a stochastic part. The deterministic aspect of the utility function $V(i)$ can be any function but most often it takes a linear form as in the equation below.

$$V(i) = a_1 X_{1i} + a_2 X_{2i} + ... + a_n X_{ni}$$

Where,

$X_{1i}, X_{2i}, ..., X_{ni}$: variable values related to the choice of alternative A_i.

$a_1, a_2, ..., a_n$: parameters to be estimated from data.

The stochastic aspect of the utility function $e(i)$ is added to $V(i)$ to make a stochastic choice model. Here, $e(i)$ follows a certain distribution to capture the uncertainty involved with the choice. By adding $e(i)$ to $V(i)$, the stochastic utility function $U(i)$ can be represented as below.

$$U(i) = V(i) + e(i)$$

Choice models assume that a certain mode of transportation is selected when it has the biggest utility function value for the traveler. Thus, the probability of choosing alternative A_i when $U(i) = U^{(1)}(i), e(i) = e^{(1)}(i)$ becomes the following.

$$p^{(1)}(i) = p[U^{(1)}(i) > U^{(1)}(j), \forall j \neq i] = p[e^{(1)}(j) < V(i) - V(j) + e^{(1)}(i), \forall j \neq i]$$

Since the probability of having $e(i) = e^{(1)}(i)$ is $f[e(i)]d[e(i)]$ — where $f[e(i)]$ denotes Probability Density Function (PDF) of $e(i)$ and $d[e(i)]$ an infinitesimal increment of $e(i)$ — the probability of choosing alternative A_i becomes the integration over the entire range of $e(i)$. By denoting $F[e(j)]$ the Cumulative Distribution Function (CDF) of $e(j)$, this probability can be expressed as the following.

$$p(i) = \int_{e(i)} F[V(i) - V(j) + e(i), \forall j \neq i] f[e(i)] d[e(i)] \tag{41}$$

The Logit model is one of the most famous stochastic choice models that is widely used in modern transportation researches. The Logit model starts from the above general choice model and adds two assumptions.

1. The stochastic variable $e(i)$ follows Gumbel's PDF

2. $e(i)$ is independent of others, i.e. e(j), e(k), etc.

Gumbel's PDF has the functional form as expressed below.

$$f(x) = \lambda e^{-x} \cdot e^{-\lambda e^{-x}}, \quad \lambda > 0, \quad -\infty < x < +\infty$$

By integration, CDF of the Gumbel's function becomes the following.

$$F(x) = e^{-\lambda e^{-x}}$$

After some calculation after putting Gumbel's distribution into equation (41), the probability to choose alternative i among others are represented as the equation below.

$$p(s) = \frac{e^{V(s)}}{\sum_{i=1}^{z} e^{V(i)}}$$

Logit models are calibrated to a specific data set using the maximum likelihood method. The process of calibrating the Logit model depends on the type of the data. If the data set is a disaggregated type, the model is configured to generate an individual's choice probability. When the data are aggregated then a choice model represents group behaviors. In that case, it is assumed that the probability of choosing a specific transportation mode is the same for all the travelers.

When the data are disaggregated type, the probability that the r^{th} traveler chooses the s^{th} mode of transportation $p(r, s)$ is represented as below.

$$p(r, s) = \frac{e^{V(r,s)}}{\sum_{i=1}^{z} e^{V(r,i)}}$$

Using the above expression, the probability $P(n_1, n_2, \ldots, n_z)$ that n_1 travelers choose mode A_1, n_2 choose mode A_2, $\ldots$, and n_z choose mode A_z is represented as the equation below.

$$P(n_1, n_2, ..., n_z) = \frac{n!}{n_1! n_2! ... n_z!} \prod_{s=1}^{z} \prod_{r \in Q_s} p(r, s)$$

To find maximum probability, the likelihood function can be defined the same as $P(n_1, n_2, ..., n_z)$ as below.

$$L = \frac{n!}{n_1!n_2!...n_z!} \prod_{s=1}^{z} \prod_{r \in Q_s} p(r,s)$$

The following equation is obtained by taking natural log on both sides.

$$lnL = ln\frac{n!}{n_1!n_2!...n_z!} + \sum_{s=1}^{z} \sum_{r \in Q_s} lnp(r,s)$$

Since the first term in the equation above is a constant, finding a condition for the maximum of L is equivalent to finding a condition for the maximum of the second term only. By defining $\overline{L} = \sum_{s=1}^{z} \sum_{r \in Q_s} lnp(r,s)$, the condition can be obtained at maximum $\overline{L}$.

If linear deterministic utility function is assumed, i.e. $V(r,s) = a_{1s}X_{1rs} + a_{2s}X_{2rs} + ... + a_{ms}X_{mrs}$, the above equation can be expanded to become the following.

$$\begin{aligned} \overline{L} &= \sum_{s=1}^{z} \sum_{r \in Q_s} lnp(r,s) = \sum_{s=1}^{z} \sum_{r \in Q_s} ln\frac{e^{V(r,s)}}{\sum_{t=1}^{z} e^{V(r,t)}} = \sum_{s=1}^{z} \sum_{r \in Q_s} [V(r,s) - ln\sum_{t=1}^{z} e^{V(r,t)}] \\ &= \sum_{s=1}^{z} \sum_{r \in Q_s} [\sum_{i=1}^{m} a_{is}X_{irs} - ln\sum_{t=1}^{z} e^{\sum_{i=1}^{m} a_{it}X_{irt}}] \end{aligned}$$

Where,

s: the number of alternative modes.

m: the number of variables in the choice criteria

$X_{1i}, X_{2i}, \ldots, X_{ni}$: evaluation of r_{th} passenger on s_{th} mode of transportation

The parameter values, a_{1s}, a_{2s}, $\ldots$, a_{ms} are estimated by choosing the values that generate maximum probability of having the data at hand. To find the condition, partial derivatives on $\overline{L}$ are taken and set to be zero as below.

$$\frac{\partial \overline{L}}{\partial a_{is}} = \sum_{i=1}^{m} X_{irs} - \sum_{i=1}^{m} X_{irs} \cdot \frac{e^{\sum_{i=1}^{m} a_{is} X_{irs}}}{\sum_{t=1}^{z} e^{\sum_{i=1}^{m} a_{it} X_{irt}}} = 0, \ s = 1, \ldots, z$$

The above equation can be solved by iterative methods. The problem with the disaggregated model is that the number of $p(r, s)$ to be calculated is equal to the product of the number of travelers and the number of transportation modes. The large number of calculation required can be a significant burden for large scale simulation.

If a choice model is based on aggregated data, it is assumed that the variable values $X's$ in the utility function is the same for all the individuals. In effect, the probability of choosing a specific transportation mode becomes constant. In equation form it becomes, $p(r, s) = p(s)$. It makes the likelihood function take the form below.

$$L = \frac{n!}{n_1! n_2! \ldots n_z!} \prod_{s=1}^{z} [p(s)]^{n_s}$$

The following equation is obtained by taking natural log function on both sides.

$$lnL = ln \frac{n!}{n_1! n_2! \ldots n_z!} + \sum_{s=1}^{z} n_s lnp(s)$$

By taking only the right term and defining $\overline{L}$ as below, it is required to find the condition that maximizes $\overline{L}$.

$$\overline{L} = \sum_{s=1}^{z} n_s lnp(s) = \sum_{s=1}^{z} n_s ln \frac{e^{V(s)}}{\sum_{i=1}^{z} e^{V(i)}} = \sum_{s=1}^{z} n_s e^{V(s)} - (n_1 + \ldots + n_z) ln(\sum_{i=1}^{z} e^{V(i)})$$

Parameters are calculated again by letting $\frac{\partial \overline{L}}{\partial a_{is}} = 0$.

Logit models are widely used to predict the performance of a new transport mode in a specific market. Impact of the air fare on a specific route is a popular usage of this model. It should also be noted that the aggregated models make assumption that everyone's utility function value is the same for the same mode of transportation service. As a result, the aggregated model deviates more from a real situation. As a tradeoff, groups of people that has similar evaluation ($X's$) for the alternatives can be aggregated and the values of $X's$ are set to the mean of each group.

APPENDIX B

AIRSPACE SIMULATION PROGRAMS

There have been many NAS modeling and simulation tools. In this section, the most well-known tools are reviewed.

B.1 Airspace Concept Evaluation System (ACES)

ACES is a fast-time NAS simulation environment for analysis of novel concepts in air traffic management. It is an agent-based simulation tool that flies individual aircraft for one day of activity. The development effort is being conducted by NASA Ames Research Center under the Virtual Airspace Simulation Technologies (VAST) sub-project of the Virtual Airspace Modeling and Simulation (VAMS) project. ACES covers the gate-to-gate operations and the highly coupled interactions in the NAS are modeled. It is the most widely used airspace simulation model by NASA. It is also used by the U.S. Air Transportation Joint Planning and Development Office (JPDO) to identify potential problems when NAS is not expanded to meet the future demand and to evaluate relevant strategies (Roth and Miraflor, 2004). ACES does not capture service providers' behavior. Therefore, demand is a static input which is not altered by the actions of the service providers.

The ACES architecture integrates an agent-based modeling framework with the distributed simulation architecture, called High Level Architecture (HLA). Figure 104 shows ACES Agents and Messages structure. Some validation efforts have been tried by Zelinski (2005).

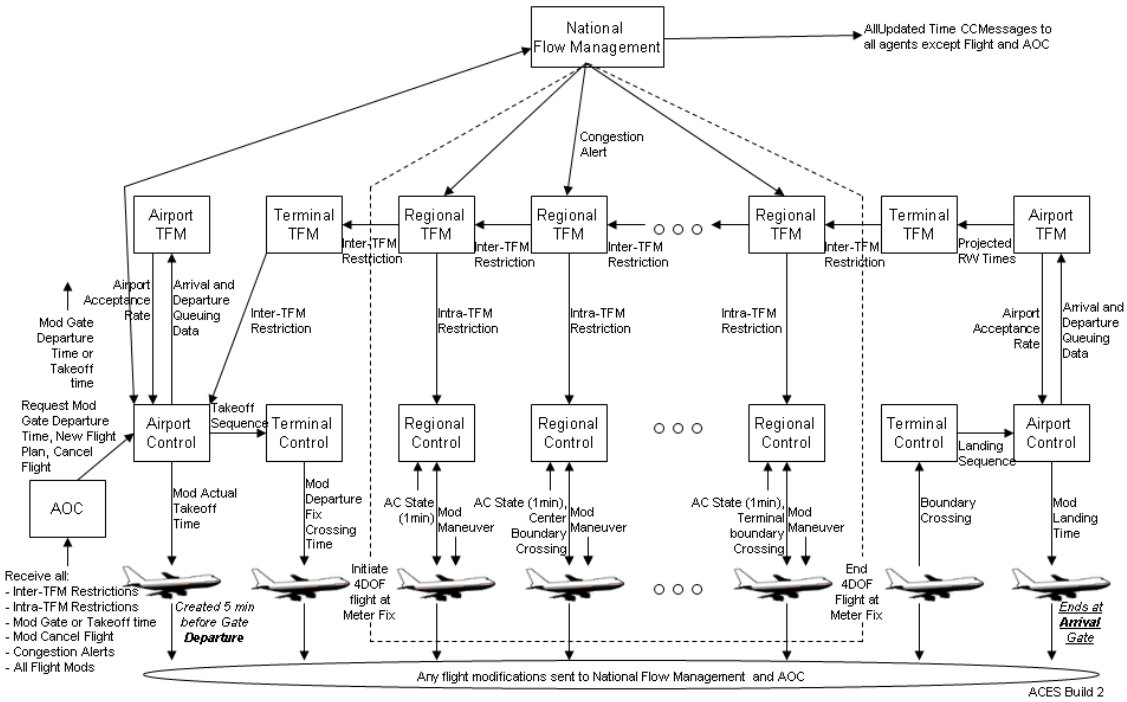


Figure 104: ACES agents and messages structure. [Source: Couluris et al. (2003)]

B.2 MIT Extensible Air Network Simulation (MEANS)

The MIT Extensible Air Network Simulation (MEANS) is a high-level event-based NAS simulation model that can be used to simulate the aircraft movements both on the ground and in the air with probabilistic capabilities. The development of MEANS started in 2001 as a tool to assess the congestion problem of an airline at a hub. Its main objectives are to study new operational concepts and developing strategic planning environment with congestion prediction. In the current version, the NAS is represented by a network of 205 airports (Clarke et al., 2007).

MEANS is a result of compromise between fidelity and tractability. This tradeoff was necessary since the NAS is a very large complex system and modeling its entirety is too time consuming and is not desired when many scenarios need to be studied in a strategic assessment. Fine-grained modules can be used for sensitive components and coarse-grained modules for insensitive ones. Therefore, many components have

simplified abstraction of the NAS. For example, in the en route phase modeling, aircrafts' tracks and conflicts resolution are not modeled and in the in-ground phase taxiways and paths are not modeled.

MEANS was implemented in C++ and its modular and extendable architecture makes it possible to combine various fidelity tools or incorporate custom made modules without changes to the core interfaces. The MEANS has seven modules in the current version as depicted in Figure 105. Aircrafts change states following the six states shown in the dotted boxes in the figure.

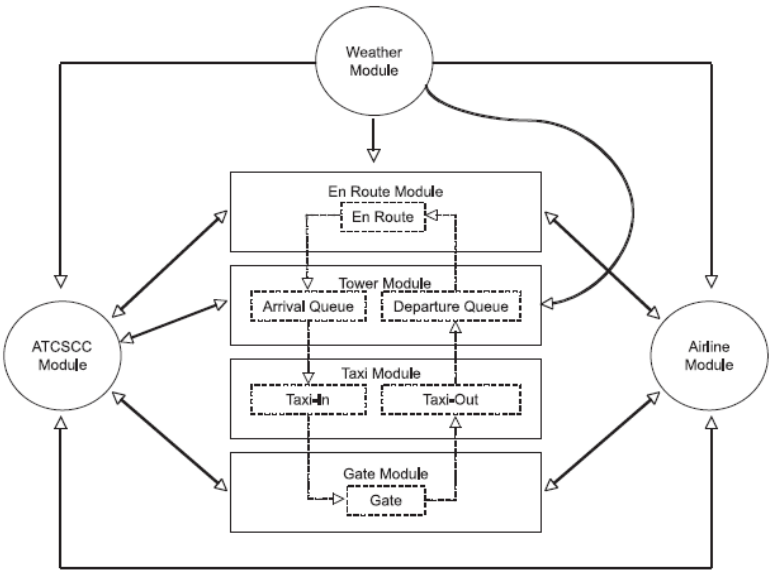


Figure 105: MEANS module relationships. [Source: Clarke et al. (2007)]

The four modules in the center of the diagram are state modules that take care of the movement of aircrafts (as well as crews and passengers) through states. The Air Traffic Control System Command Center (ATCSCC) module and the airline module change the flight schedule and are referred to as decision-making modules. The characteristics of the seven modules are listed in Table 23.

The inputs that MEANS needs are the aircraft to be used and the sequence of flight schedules (origin, destination, scheduled departure time, and scheduled arrival time) that each aircraft are to fly. This flight data are determined using the Airline Service

Table 23: Characteristics of seven modules in MEANS

State Modules	Gate Taxi Tower Enroute
Informational Module	Weather
Decision Making Module	Air Traffic Control Systems Command Center Airline

Quality Performance (ASQP) database. Since the ASQP database only contains data provided by airlines that account for at least one percent of all domestic passengers, artificial flight legs are added to match the hourly real flight demand represented in the Enhanced Traffic Management System (ETMS) database. The flights derived from the ASQP database are tracked but the artificial flights plays a role to fill the demand in the system, such as the departure queue. MEANS requires different levels of information for its modules. Table 24 summarizes the data required.

Table 24: Data requirements for MEANS modules

	Data Requirements	Data Source
Schedule	Detailed scheduled flight information	ASQP database ETMS database
Airport Capacities	Runway configuration and capacity	FAA Benchmark Report Theoretical Generation
Airborne, Taxi, Ground Times	Historical distributions	ASQP
Weather	Severity	NOAA weather record

Taxi-out and taxi-in times can be modeled as a single value, historical distribution, or estimation based on aircraft passing behavior. In the simulation, the Ground Delay Program (GDP) is initiated by the ATCSCC module when predicted demand is higher than the predicted capacity by a specified amount for a specified duration at an airport. GDP initiates ground stops for all flights in the NAS that are scheduled to land at the airport during the period when the predicted demand exceeds the predicted capacity. The airline module cancels or reschedules flights in response to

delays. In the simulation, each passenger is tracked. When flights are cancelled or delayed such that the connection is not possible for their initial itinerary, they are re-accommodated on later flights when possible.

The primary output from the MEANS is a set of files containing flight and passenger information. Its output includes detailed time information for each aircraft, desired and obtained itineraries for passengers, and delay and cancellation characteristics. Its stochastic capabilities are achieved through multiple runs with the same data and different random seed.

B.3 LMINET

LMINET is a network queueing model of the NAS developed by Logistics Management Institute (LMI) for NASA. It is a fast-time model that covers flight schedules of over 600 Official Airline Guide (OAG)-derived airports, over 3000 National Plan of Integrated Airport Systems (NPIAS)-derived airports, and OAG-derived international airports. It can calculate delays at 102 airports and en route airspace constraints at 995 sectors. The inputs to the simulation include the flight schedule, flight trajectories, airport runway configuration, and airport weather conditions.

LMINET is composed of 64 major US airports and they account for more than 80 percent of the airline flights in 1997. Airports, TRACONs, and ARTCC sectors are connected with sequences of queues. LMINET is driven by demand and weather condition. Both of them are provided as hour-by-hour values and aggregate level analysis is performed. It is not a simulation but an analytical model that does not track aircraft. Queues are represented as arrival and service processes with distributions. The model is stochastic in nature. The LMINNET is driven by the scheduled departures from the airports within the network and the scheduled arrivals from the airports outside the network. Operations at airports are modeled by a queueing network shown in Figure 106.

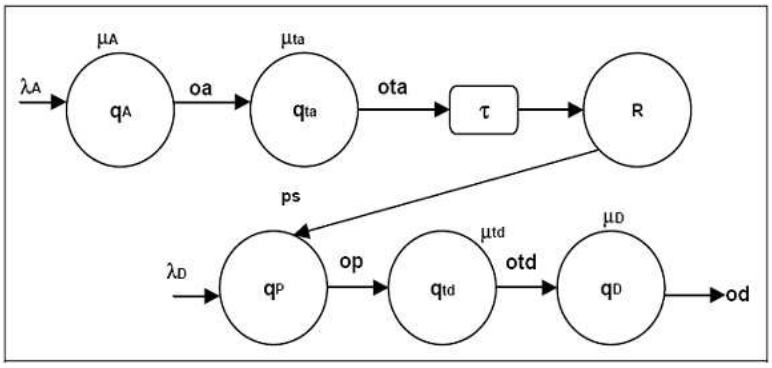


Figure 106: Queues in the LMINET airport model. [Source: Long, Lee, Johnson, Gaier and Kostiuk (1999)]

The airport queueing network model proceeds as follows.

1. Traffic enters the arrival queue, qA, with a Poisson process with parameter $\lambda A(t)$

2. The arrival queue has a Poisson service time with parameter $\mu A(t)$

3. Traffic enters the taxi-in queue, qta, that has a Poisson service with parameter μta

4. After turnaround delay(τ), the traffic enters the ready-to-depart reservoir, R

5. Departures enter the queue for aircraft, qp, with a Poisson process with parameter $\lambda D(t)$

6. Departure aircraft are assigned with a Poisson process with parameter $\mu p(t)$ and R is reduced by 1. After the assignment, the aircraft enters the taxi-out queue, qtd, that has a Poisson service time with parameter $\mu td(t)$

7. The aircraft enters the departure runway queue, qD, that has a Poisson service process with parameter $\mu D(t)$

8. The aircraft is fed into the rest of LMINET.

ARTCC and TRACON sectors are modeled as M/Ek/N/N+q multi-server queues, which are Poisson arrivals, service time of Erlang distribution with parameter k, N servers, and maximum number in the system to be N+q. TRACON of an airport is modeled with two arrival sectors and one departure sector.

B.4 Simulation Model (SIMMOD)

SIMMOD has a highly detailed network representation of airfield and airspaces. Complex network models need to be developed by the user. Input requirements include flight paths and paths between gates and runways. SIMMOD has a database with 19 types of aircraft performance characteristics. SIMMOD generates detailed flight level outputs including aircraft travel times, delays, fuel consumption, and flow information. It can simulate probabilistic events and produces detailed flight level outputs. SIMMOD is frequently used in airport surface operations and capacity studies. Traffic follows a pre-specified network of nodes and links with the pre-specified operation strategies. It is a one-dimensional model that checks conflicts only along the longitudinal path. Due to its non-modular architecture, SIMMOD is not a proper tool for radical ATM concept studies.

A single run of 24 hour simulation at a major airport takes about 3-5 minutes. SIMMOD is a widely used airport and airspace model in the world. Its major application has been in the study of the impacts of various operational concepts at airports and reconfiguration of regional or terminal airspace on the capacity and delay. SIMMOD requires significant user training and understanding of ATM and airport operations.

B.5 The National Airspace System Performance Capability (NASPAC)

NASPAC is a macroscopic system-wide model of air traffic flows and delays that was initially developed by the MITRE Corporation for the FAA in the late 1980s. It is

a useful tool when assessing new concepts in traffic flow management. It is the first model that was explicitly developed for the purpose of assessing delay propagation and congestion in the NAS. The simulation tracks aircrafts through the daily itineraries and generates deterministic information on delays and traffic flows for the input. It simulates individual flight but the output is aggregate delays and flow rates. The simulation requires complete schedule of aircraft itineraries in the interested airspace, airport capacities and other ATM resources, and aircraft performance as inputs. The modeling of ATM resources is not detailed. The level of details of the modeling can be varied for the problem at hand. NASPAC needs significant user training and input data.

B.6 Total Airspace and Airport Modeler (TAAM)

TAAM is a detailed large-scale network simulation model of airport and airspace developed by The Preston Group (TPG) and Australian Civil Aviation Authority (CAA). TAAM can simulate most ATM operations and the entire gate to gate processes are modeled in detail. TAAM is a 4D flight simulation that is capable of entire air traffic system simulation. TAAM requires significant input data that represent the entire air traffic system. The inputs includes airport descriptions, airspace route and sector layouts, geographical features, air traffic control and airport usage rules, aircraft performance characteristics, etc. To use TAAM, significant user training and input data preparation are required.

B.7 Detailed Policy Assessment Tool (DPAT)

DPAT is a fast-time discrete-event simulation whose scope covers global air traffic. DPAT is built on a parallel discrete-event simulation engine, called Georgia Tech Time Warp (GTW), which uses optimistic computing technology and achieves very fast run times. It is a very fast network of queueing model that tracks individual flights and their itineraries. DPAT predicts air traffic delays at airports and airspace sectors and

propagates delays over the flights' itineraries. DPAT is derived from National Airspace System Performance Analysis Capability (NASPAC) but it is much faster due to parallel processing and improvements in coding. It has significant data requirements and much of the required data should be generated using other simulations

B.8 Future ATM Concepts Evaluation Tool (FACET)

FACET is NAS simulation that is designed for the evaluation of advanced ATM strategies. It is written in "Java" and "C" and it is platform independent. Its' modular architecture provides flexibility for fast reconfiguration for new ATM concepts. It mainly focuses on airspace operations and lacks in airport operations modeling. FACET models enroute airspace over the entire continental U.S. It has 4-D trajectory modeling capabilities, dynamic models for turns and acceleration/deceleration, climb/descent performance models for 66 aircraft types that are mapped to over 500 types, and weather models. Figure 107 shows schematic overview of FACET. FACET interfaces with ETMS data. It currently runs on Sun, SGI, PC, and Mac.

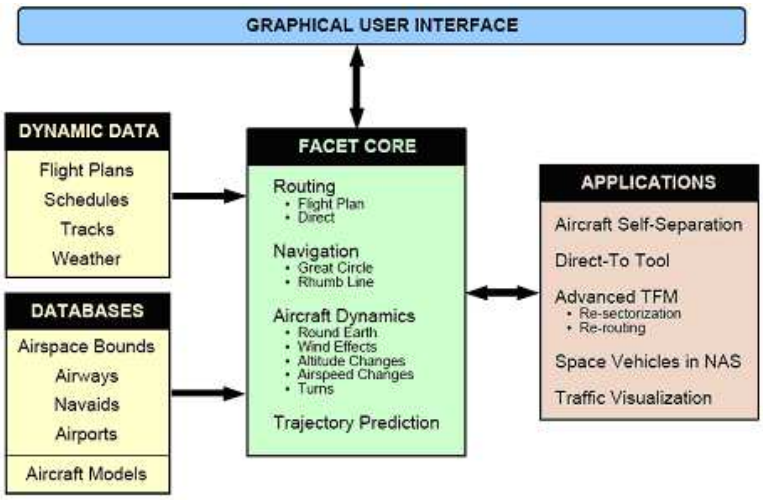


Figure 107: Schematic overview of FACET. [Source: National Aeronautics and Space Administration (2003)]

APPENDIX C

DELAY PROBLEMS OF THE NAS

C.1 Delay problems are getting worse

With the growth of enplanements and passenger miles, flight demands are also expected to grow proportionally. When flight demands reach the capacity of busy hub airports, flight and passenger delay problems become focal issues. Figure 108 shows the flight operations, on-time, arrival delayed, canceled and diverted flights from 1988 to 2005, taken from the Bureau of Transportation Statistics (BTS)'s Airline On-Time Performance database (Bureau of Transportation Statistics, 2009*b*). Airline On-Time Performance data contains data reported by US certified air carriers and accounts for at least one percent of the domestic scheduled passenger revenues. The delay criteria used in the database is when a flight arrives more than 15 mins after its published arrival time. Definitions of some of the frequently used BTS terms are listed in Table 25.

Seen in the Figure 108, arrival delays have been increasing since 2002 as the air travel demand recovered from the shock of the 2001 terrorist attack. As the congestion and delay worsened, customer complaints also increased correspondingly (Department of Transportation, 2009), resulting in additional cost to both the airlines and air transportation users. As such many researches and modeling studies were dedicated to the delay analysis and improvement efforts of these delay problems (Viken et al., 2006*a*; Department of Transportation et al., 2004).

The Federal Aviation Administration (FAA) has developed airport capacity benchmarks for 31 of the nation's busiest airports in 2001 and updated it in 2004 expanding

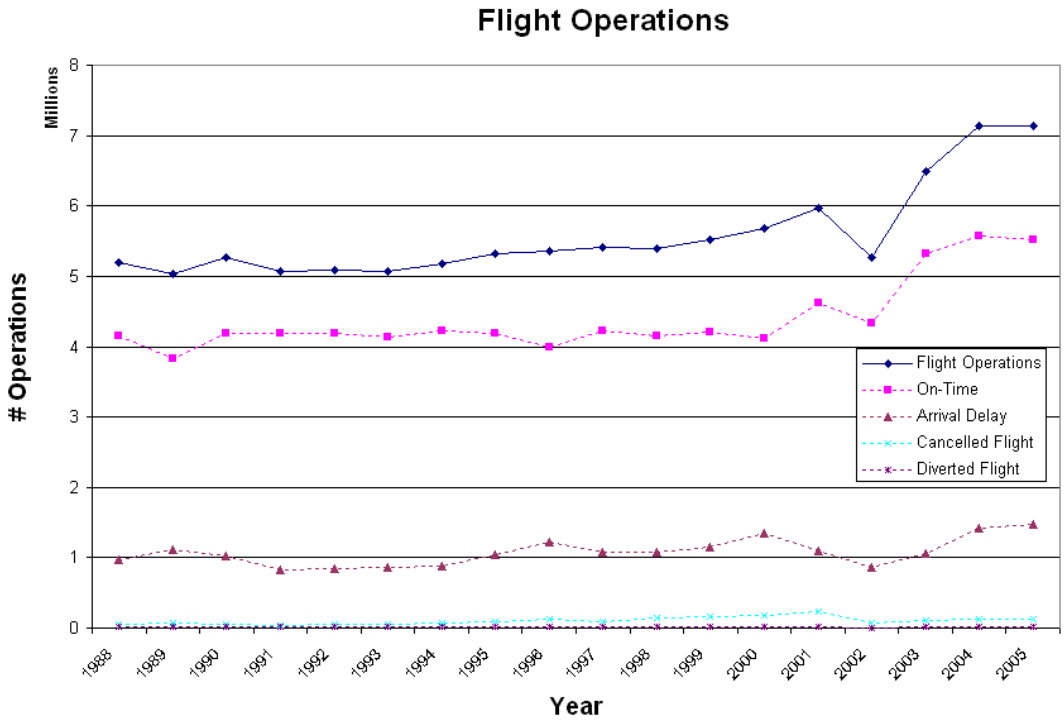


Figure 108: Yearly on-time characteristics

Table 25: Flight related terms and definitions. [Source: BTS website]

Terms	Description
Arrival Delay	Arrival delay equals the difference of the actual arrival time minus the scheduled arrival time. A flight is considered on-time when it arrives less than 15 minutes after its published arrival time.
Departure Delay	The difference between the scheduled departure time and the actual departure time from the origin airport gate.
In-Flight Time	The total time an aircraft is in the air between an origin-destination airport pair, i.e. from wheels-off at the origin airport to wheels-down at the destination airport.
Late Flight	A flight arriving or departing 15 minutes or more after the scheduled time.
Taxi-In Time	The time elapsed between wheels down and arrival at the destination airport gate.
Taxi-Out Time	The time elapsed between departure from the origin airport gate and wheels off.

its analysis to a total of 35 airports (Federal Aviation Administration, 2004). It addresses the relation between airline demand and airport runway capacity. When the scheduled flight demand at an airport exceeds its capacity at the moment, resulting delays spread out to the affected routes. Generally, the delay problem is most severe in Summer when flight demand and weather disturbances are highest. Airport Capacity Benchmark Report 2004 provides capacity benchmarks for 35 of the nation's busiest airports (See Table 26.). These 35 airports are closely monitored and are part of Operational Evolution Partnership (OEP) (Federal Aviation Administration, 2007). The benchmark numbers in Table 26 show the maximum number of flight operations that an airport can routinely handle in an hour based on typical runway usage.

Delays also happen as a consequence of fluctuations in demand. Even when the average demand is below average capacity, delays can be severe if fluctuation is high. Figure 109 shows capacity and demand at Chicago - O'Hare International Airport (ORD) on July12, 2004, illustrating demand exceeding the capacity in regular days due to fluctuations.

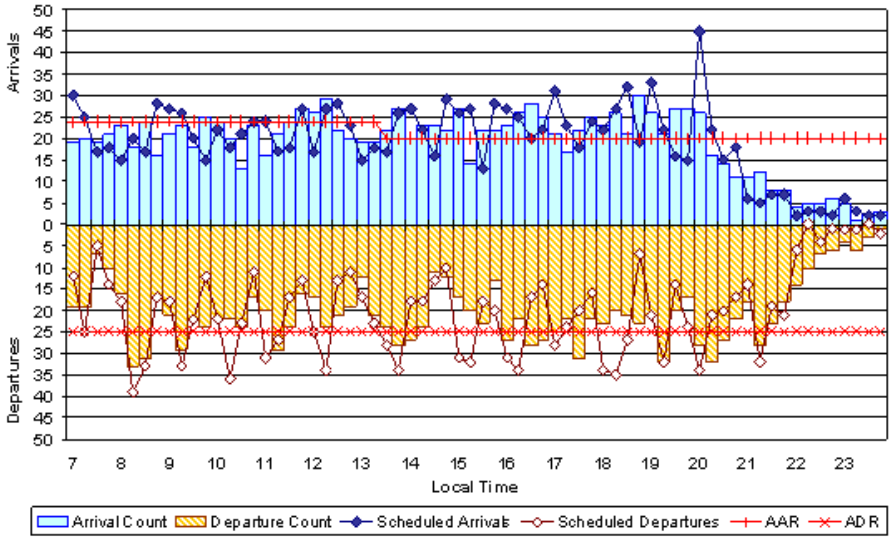


Figure 109: Capacity and demand at ORD on July 12, 2004 (Monday). [Source: Federal Aviation Administration (2004)]

Table 26: Capacity benchmarks at OEP 35 airports, 2004. [Source: Federal Aviation Administration (2004)]

Airport		Optimum	Marginal	IFR
ATL	Atlanta Hartsfield-Jackson International	180-188	172-174	158-162
BOS	Boston Logan International	123-131	112-117	90-93
BWI	Baltimore-Washington International	106-120	80-93	60-71
CLE	Cleveland Hopkins	80-80	72-77	64-64
CLT	Charlotte/Douglas International	130-131	125-131	102-110
CVG	Cincinnati/Northern Kentucky International	120-125	120-124	102-120
DCA	Ronald Reagan Washington National	72-87	60-84	48-70
DEN	Denver International	210-219	186-202	159-162
DFW	Dallas/Fort Worth International	270-279	231-252	186-193
DTW	Detroit Metro Wayne County	184-189	168-173	136-145
EWR	Newark Liberty International	84-92	80-81	61-66
FLL	Fort Lauderdale-Hollywood International	60-62	60-61	52-56
HNL	Honolulu International	110-120	60-85	58-60
IAD	Washington Dulles International	135-135	114-120	105-113
IAH	Houston George Bush Intercontinental	120-143	120-141	108-112
JFK	New York John F. Kennedy International	75-87	75-87	64-67
LAS	Las Vegas McCarran International	102-113	77-82	70-70
LAX	Los Angeles International	137-148	126-132	117-124
LGA	New York LaGuardia	78-85	74-84	69-74
MCO	Orlando International	144-164	132-144	104-117
MDW	Chicago Midway	64-65	64-65	61-64
MEM	Memphis International	148-181	140-167	120-132
MIA	Miami International	116-121	104-118	92-96
MSP	Minneapolis-St Paul International	114-120	112-115	112-114
ORD	Chicago O'Hare International	190-200	190-200	136-144
PDX	Portland International	116-120	79-80	77-80
PHL	Philadelphia International	104-116	96-102	96-96
PHX	Phoenix Sky Harbor International	128-150	108-118	108-118
PIT	Greater Pittsburgh International	152-160	143-150	119-150
SAN	San Diego International - Lindbergh Field	56-58	56-58	48-50
SEA	Seattle-Tacoma International	80-84	74-76	57-60
SFO	San Francisco International	105-110	81-93	68-72
SLC	Salt Lake City International	130-131	110-120	110-113
STL	Lambert-St. Louis International	104-113	91-96	64-70
TPA	Tampa International	102-105	90-95	74-75

Aviation gridlock experienced on the NAS calls for a system-wide view to understand the situation and seek solutions (Transportation Research Board, 2001). In this light, the Joint Planning and Development Office (JPDO) is developing a Concept of Operations (CONOPS) for the Next Generation Air Transportation System (NextGen) (Joint Planning and Development Office, 2007). FAA developed various methods to measure the performance of the NAS (Federal Aviation Administration, 2001). O'Connor and Rutishauser (2001) studied the dynamic reductions in aircraft separation using the Aircraft VOrtex Spacing System (AVOSS). In the AVOSS, the separation criteria applied to aircraft for wake vortex avoidance is selected for improvements. Acknowledging the fact that wake hazard durations are substantially reduced in many ambient conditions, the overly conservative separation criteria was relaxed in a real-time proof-of-concept field study at the Dallas Ft. Worth International Airport in July, 2000. AVOSS used weather sensors, wake sensors, and analytical wake prediction algorithms to obtaine 6% of the airport throughput gains, with peak values close to the theoretical maximum of 16%. The average gain in airport delay when applied to major airports were 15-50%.

Airport delay can be amplified by small disturbances. For example, consider the situation where airplanes arrive at the airport at a rate of one per minute. Assuming that 10 airplanes are already waiting for landing and the airport capacity for landing is one per minute, then the delay time for a newly arriving airplane is 10 minutes. If a slight disturbance caused by weather condition delays the landing for 5 minutes, the queue would now grow to 15 airplanes waiting for landing. It should be noted that a small amount of delay at an airport in H&S network can cause a significant propagation throughout the network.

C.1.1 Delay causes

There are mainly four causes major routine delay.

- Over-scheduling exceeding the capacity of airports: The delay impact of over-scheduling can be significant due to queueing effect and last throughout the day.

- Increased number of regional jets generate large demand on airport and airspace: Scheduled flights aboard regional jets increased from 9 percent in May 2000 to 29 percent in May 2004. This was due to the airliners relying on their regional partners to serve smaller markets (Department of Transportation, 2004).

- Increased low-cost carriers at large and medium hubs: Low-cost carriers serving small markets are expanding to large and medium markets. As a result of the low airfares provided by both the low-cost carriers and the response of major carriers, new passenger demand is created. This new demand can generate significant delays on affected airports and airspaces. Low-cost carriers' market share have grown from 15 percent in 2000 to 21 percent in 2004 (Department of Transportation, 2004).

- Increased level of security screening after $September11^{th}$: After the terrorist attack in September 11, 2001, the Transportation Security Administration (TSA) was created. By the end of 2002, TSA deployed a federal work force for screening passengers and baggages. The queues for security check point at busy airports routinely extends an hour or more.

Delays in the NAS is mainly a result from air traffic density and weather conditions. Occasional delay causes include aircraft mechanical problems.

C.1.2 Alleviating air transportation network delays

Short-term strategies for alleviating air transportation delays:

- Airlines' voluntary schedule changes.

- Allocation of capacity administrative intervention to reduce or change scheduled flights for temporary delay reduction: This approach might be sought for a busy airport where air carriers are reluctant to adjust schedules for fear of losing market share. This was the approach undertaken by FAA to alleviate delay problems at the Chicago O'Hare airport.

Long-term strategies for alleviating air transportation delays:

- New runway construction: New runways provide the most capacity benefit to the air transportation network. Seven new runways were constructed between fiscal year 2001 and 2005, four runways were scheduled to open in fiscal year 2006, with six more under construction (Federal Aviation Administration, 2007).

- Advanced operational concept: Dynamic airspace configuration, airspace super-density operations, coordinated arrival/departure, etc can also increase capacity and reduce delays (Joint Planning and Development Office, 2007).

To increase the capacity for the the demand increase, there are eight ongoing runway projects in OEP (Federal Aviation Administration, 2007). Even though the additional runways are a guaranteed way of increasing capacity, it is not always an option due to space and financial constraints. Therefore, it is important to identify options and seek the best solutions.

C.2 Schedule padding as a response to congestion at hubs

Time information recorded in the Airline On-Time Performance Data are divided into taxi-out, airtime, and taxi-in time. When the data between 1995 and 2005 for ATL $\rightarrow$ ORD was plotted as a distribution shown in Figure 110, characteristics of a shifted and flattened distribution was observed. This flattened and shifted distribution can be interpreted as it took more time to travel from Atlanta to Chicago on average.

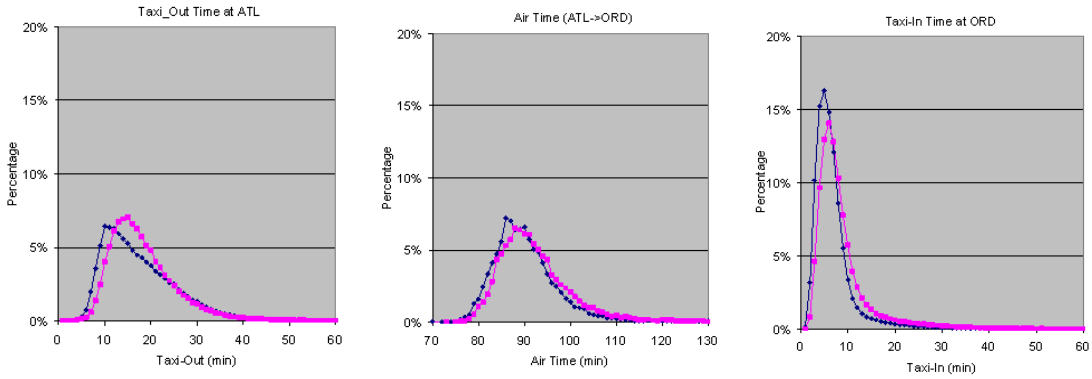


Figure 110: Changes in taxi-out, taxi-in, and air time for flights from ATL to ORD in 1995 (blue) and 2005 (red) (* Data greater than 60 mins are excluded for taxi-in and taxi-out)

It is also interesting to see a similar trend in the total arrival delays on the NAS. Figure 111 shows a crude assessment of the arrival delay distribution at Hartsfield-Jackson Atlanta International Airport (ATL) with respect to the baseline 1988 data.

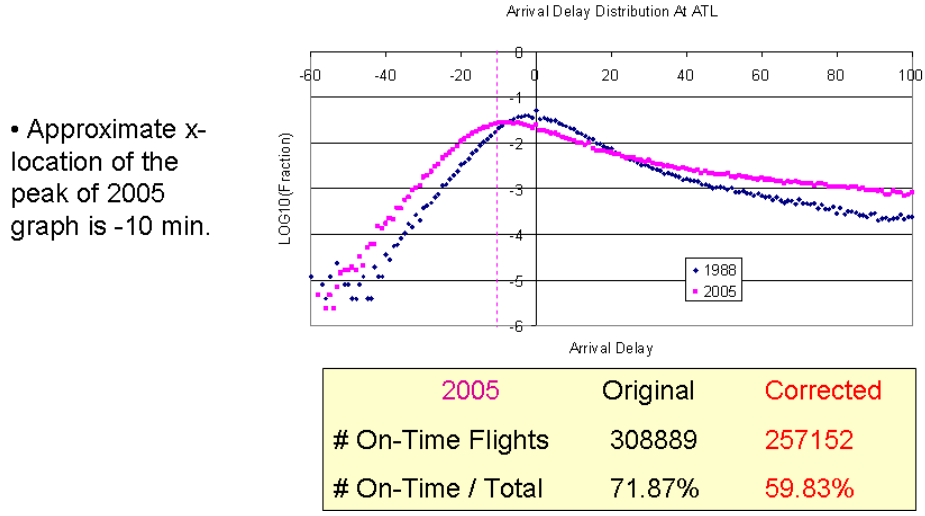


Figure 111: Arrival delay trend change at ATL

In this study, the highest point on each graph was assumed to be the airlines' targeted scheduled delay. By shifting the line for 2005 to zero, the on-time percentage reduced from 72 percent to 60 percent. This crude assessment shows that the delay

problem is actually worse than it appears. Congestion and resulting delays led airlines to add slack time to the flight schedule. As an example, the average block time increase on the route from ATL (Atlanta,GA) to DFW (Dallas/Ft.Worth,TX) is shown in Figure 112. This shows that scheduled block time increased simultaneously with actual block time. A possible interpretation of this phenomenon is that as the delay problem worsens and customer complaints increases, the airlines increased published block time on the travel routes between busy airports. In this case, average block time is interpreted as a base flight time for the segment.

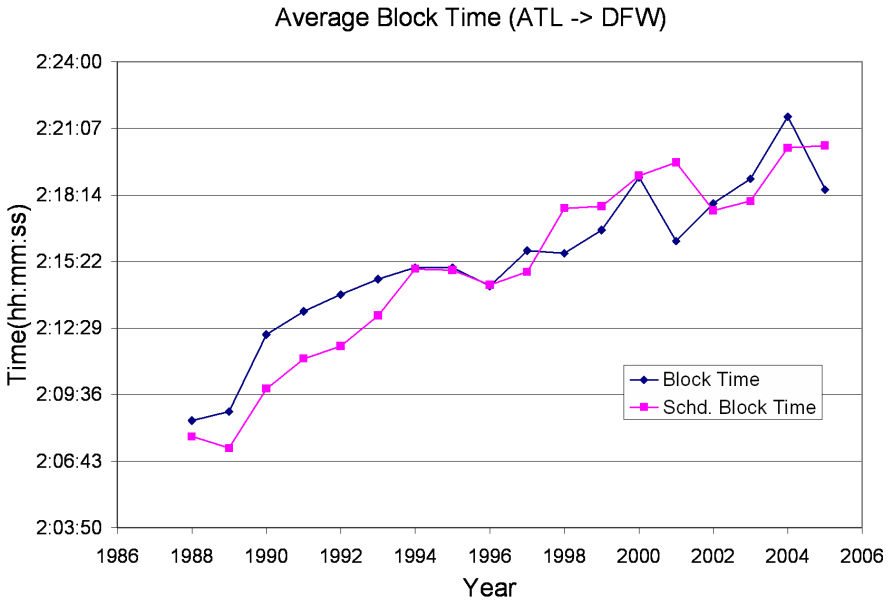
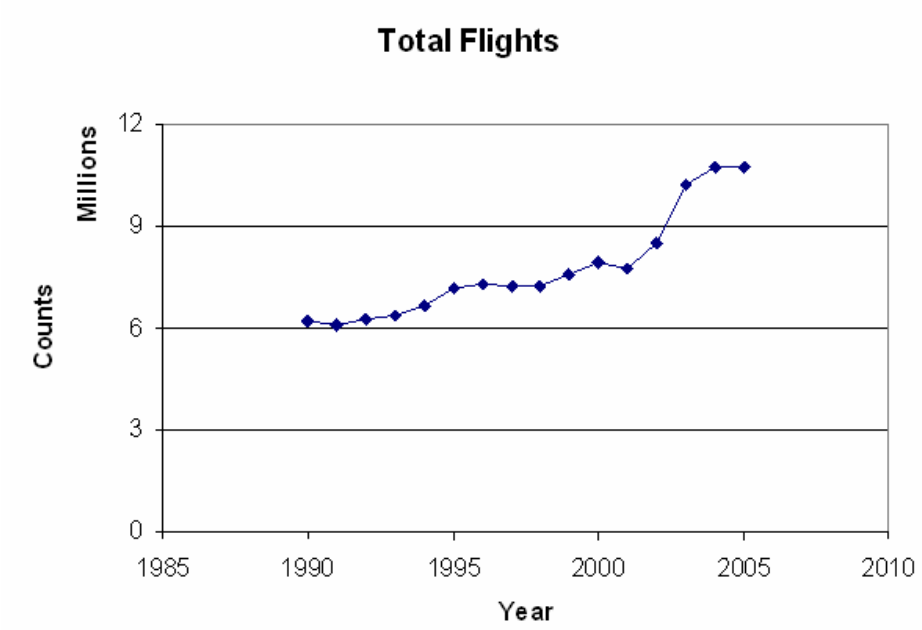


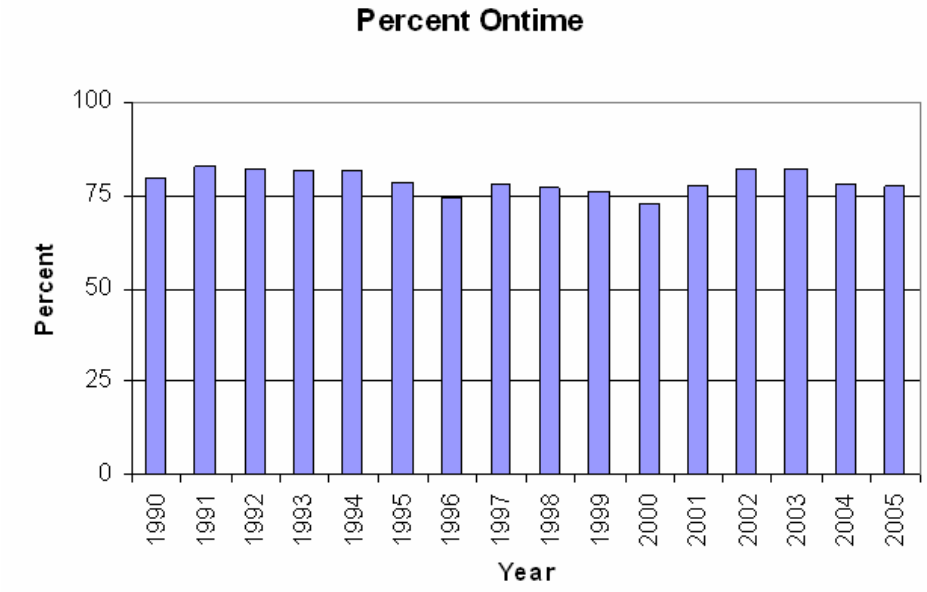
Figure 112: Block time change: ATL → DFW (* Data with unusual values ($\leq 1:10:00$ or $\geq 5:00:00$) are excluded)

This also accounts for the fact that published delay percentage does not grow even though the general public experiences more delays. By shifting the delay distribution through schedule padding, airlines can maintain on-time arrival statistics despite high demand. Figure 113 shows the total flight operations and percent on-time characteristics.

As block time padding is an airlines response to the degraded NAS performance,



(a) Total flight counts. Source: T100D segment data



(b) Percent on-time. Source: On-time performance data

Figure 113: Total flights and percent on-time

this adaptive change is also considered by including a reference point for delay measure. Figure 114 shows an example of the effect of block time padding consideration. In this case study, flight delays from ATL to DFW were corrected with respect to the baseline block time in 1988. It clearly shows correlation between the flight operations and on-time performance.

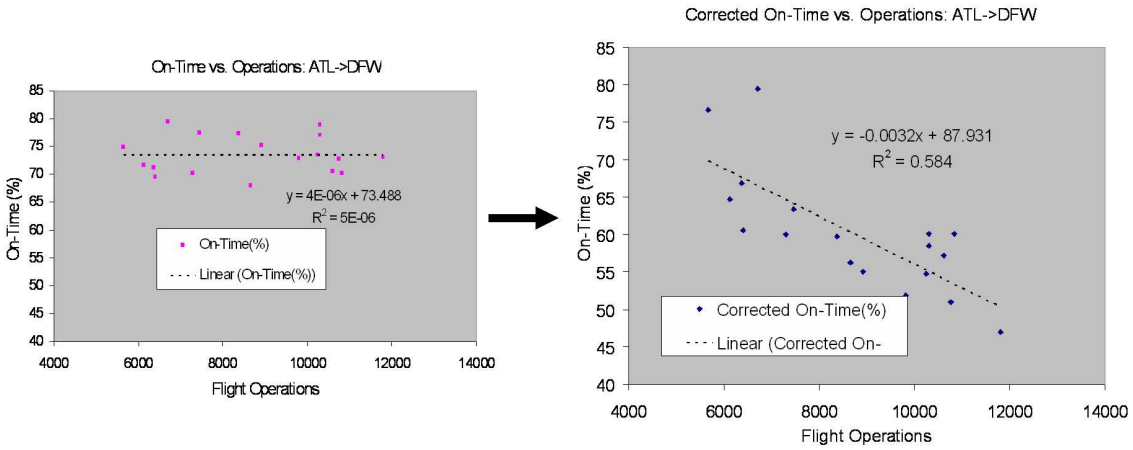


Figure 114: Correlation exists with corrected data: Case study: ATL → DFW

The corrected on-time trend calculation for this case was conducted with the following procedure.

1. Baseline scheduled block time was based on 1988 average.

2. Difference between baseline and current scheduled block time (padding) was subtracted from each scheduled arrival time to obtain the corrected scheduled arrival (See Table 27 for padding calculation) time.

   (Corrected Scheduled Arrival Time = Scheduled Arrival Time − Padding)

3. The new delay criteria based on this new scheduled arrival time.

   (Corrected Arrival Delay = Actual Arrival Time−Corrected Scheduled Arrival Time)

Table 27: Average scheduled blocktime and padding

Year	Schd.	Padding
1988	2:07:31	0:00:00
1989	2:06:58	-0:00:33
1990	2:09:25	0:01:54
1991	2:10:33	0:03:01
1992	2:11:20	0:03:49
1993	2:12:56	0:05:24
1994	2:14:40	0:07:09
1995	2:14:36	0:07:04
1996	2:13:55	0:06:24
1997	2:14:33	0:07:02
1998	2:17:08	0:09:36
1999	2:17:19	0:09:48
2000	2:18:40	0:11:09
2001	2:19:13	0:11:41
2002	2:17:33	0:10:02
2003	2:17:55	0:10:24
2004	2:20:14	0:12:43
2005	2:20:12	0:12:40

APPENDIX D

NAS-RELATED DATA

D.1 Airline On-Time Performance data

The airline on-time performance data is offered by the Bureau of Transportation Statistics (BTS). Title 14, Code of Federal Regulations (CFR), part 234 - airline service quality performance reports requires domestic certified air carriers that account for at least 1 percent of domestic scheduled passenger revenues to report on-time performance data. The reporting carriers and the number of operations in descending order is in Figure 115.

	CARRIER	Year 1988	CARRIER	Year 1989	CARRIER	Year 1990	CARRIER	Year 1991	CARRIER	Year 1992	CARRIER	Year 1993	CARRIER	Year 1994	CARRIER	Year 1995	CARRIER	Year 1996
1	DL	750,228	DL	783,320	US	1,002,485	US	907,184	DL	916,593	DL	898,896	DL	874,526	DL	884,019	DL	888,306
2	AA	693,000	AA	723,252	DL	824,062	DL	874,791	US	886,305	US	833,131	US	857,906	US	778,835	WN	757,419
3	UA	587,144	US	712,342	AA	712,060	AA	725,191	AA	782,371	AA	786,696	AA	722,277	UA	724,807	UA	735,266
4	PI	470,957	UA	574,674	UA	606,713	UA	630,093	UA	639,349	UA	649,086	UA	638,750	WN	693,101	US	728,534
5	CO	457,031	NW	451,752	NW	461,745	NW	460,230	NW	474,931	WN	488,332	WN	565,426	AA	688,471	AA	655,539
6	NW	431,440	CO	410,322	CO	427,570	CO	424,734	WN	422,912	NW	474,831	CO	484,834	NW	521,693	NW	543,041
7	EA	388,187	WN	288,802	WN	322,759	WN	368,734	CO	421,974	CO	419,560	NW	482,798	CO	434,425	CO	414,410
8	AL	361,059	PI	286,518	TW	264,182	TW	231,769	TW	246,449	TW	244,782	TW	258,205	TW	272,452	TW	282,176
9	TW	275,819	TW	276,471	EA	252,013	HP	219,896	HP	199,066	HP	176,756	HP	177,851	HP	194,508	HP	203,001
10	WN	262,422	HP	195,588	HP	216,314	AS	98,952	AS	102,207	AS	98,431	AS	117,475	AS	135,124	AS	144,291
11	HP	180,871	EA	169,704	AS	97,934	ML	70,622										
12	US	133,324	AS	89,122	PA	83,056	PA	64,729										
13	AS	89,822	PA	79,333														
14	PA	72,264																
15	PS	41,911																

	CARRIER	Year 1997	CARRIER	Year 1998	CARRIER	Year 1999	CARRIER	Year 2000	CARRIER	Year 2001	CARRIER	Year 2002	CARRIER	Year 2003	CARRIER	Year 2004	CARRIER	Year 2005
1	DL	921,850	DL	915,095	DL	914,130	WN	911,699	WN	957,145	WN	956,745	WN	958,566	WN	990,404	WN	1,036,034
2	WN	794,849	WN	815,069	WN	853,683	DL	908,029	DL	835,236	AA	852,439	AA	752,241	AA	698,548	AA	673,569
3	UA	743,847	UA	748,459	UA	774,370	UA	776,559	AA	716,985	DL	728,758	DL	660,617	DL	687,638	DL	658,302
4	US	718,751	US	696,030	US	714,430	US	748,624	UA	704,977	UA	587,887	UA	543,957	UA	555,812	MQ	532,032
5	AA	663,954	AA	653,919	AA	692,653	AA	742,265	US	688,748	NW	513,331	NW	499,160	NW	507,286	OO	517,454
6	NW	537,152	NW	527,512	NW	537,180	NW	551,337	NW	533,045	US	511,664	MQ	429,098	MQ	483,088	UA	485,918
7	CO	403,349	CO	399,273	CO	400,014	CO	393,036	MQ	489,418	MQ	439,984	US	411,956	OO	463,448	NW	480,832
8	TW	274,037	TW	274,073	TW	275,830	TW	267,131	CO	377,281	CO	325,246	OO	396,801	US	419,919	US	425,609
9	HP	206,304	HP	203,071	HP	209,496	HP	219,160	TW	244,721	HP	196,111	RU	328,086	OH	373,346	RU	403,767
10	AS	147,750	AS	152,220	AS	156,098	AS	154,171	HP	209,979	AS	159,194	CO	302,742	RU	367,703	OH	381,337
11							AQ	11,036	AS	156,447			DH	291,600	CO	299,329	EV	304,647
12									AQ	53,798			EV	273,712	EV	278,861	CO	296,661
13													HP	189,519	DH	264,955	HP	197,853
14													AS	161,594	HP	195,939	FL	195,392
15													FL	144,700	AS	166,992	AS	158,100
16													TZ	69,176	FL	162,558	DH	136,492
17													B6	67,184	B6	89,834	B6	111,050
18													HA	7,831	TZ	75,533	F9	53,255
19															HA	48,077	HA	48,183
20																	TZ	44,109

Figure 115: Reporting carriers of on-time performance data in descending order

The raw files of on-time performance data are available for download from BTS web site (`http://www.transtats.bts.gov`). The data range is from 1987 and it is updated monthly. A snap shot of the download web site for the airline on-time performance data is shown in Figure 116.

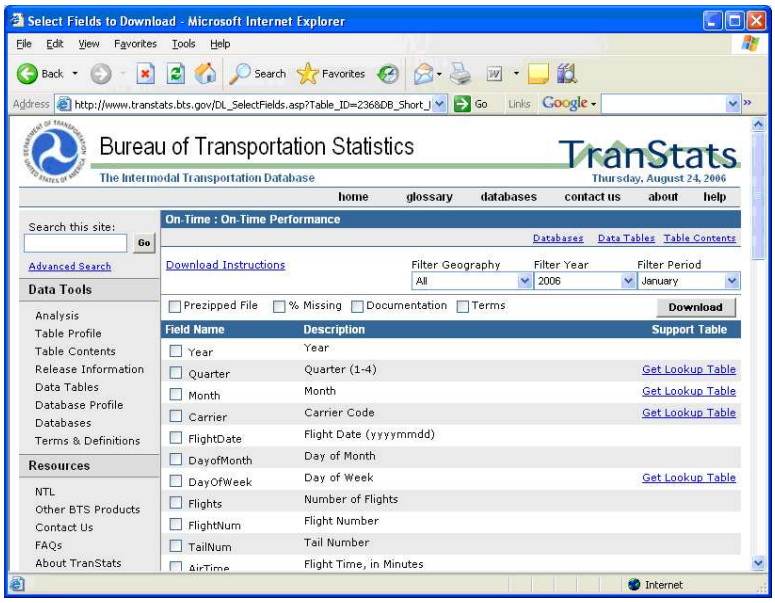


Figure 116: Snapshot - Airline on-time performance data download site

Various arrival delay information can directly be obtained from the airline on-time performance data as shown in Table 28.

Table 28: Arrival delay data. [Source: Airline On-Time Performance Data by BTS]

Year	# Operations	# Ontime	Ontime(%)	# Arrival Delays	Arrival Delays(%)	# Cancelled	Cancelled (%)	# Diverted	Diverted(%)
1988	5,195,479	4,154,387	79.96158	976,585	18.79682316	50,087	0.009640497	14,420	0.277549
1989	5,041,200	3,832,730	76.02813	1,119,466	22.20633976	74,165	0.014711775	14,839	0.294354519
1990	5,270,893	4,183,118	79.36261	1,019,363	19.33947435	52,458	0.009952393	15,954	0.302681159
1991	5,076,925	4,186,857	82.46836	833,978	16.42683317	43,505	0.008569163	12,585	0.24788627
1992	5,092,157	4,189,590	82.27535	838,347	16.46349474	52,836	0.010375957	11,384	0.223559486
1993	5,070,501	4,139,064	81.63028	861,259	16.98567854	59,845	0.011802581	10,333	0.203786569
1994	5,180,048	4,219,794	81.46245	881,408	17.01544078	66,740	0.01288405	12,106	0.233704398
1995	5,327,435	4,185,788	78.57042	1,039,250	19.50751159	91,905	0.017251266	10,492	0.196942806
1996	5,351,983	3,989,281	74.53837	1,220,045	22.79612996	128,536	0.024016519	14,121	0.26384613
1997	5,411,843	4,218,165	77.94323	1,083,834	20.02707765	97,763	0.018064641	12,081	0.22323264
1998	5,384,721	4,156,980	77.19954	1,070,071	19.87235736	144,509	0.026836859	13,161	0.244413777
1999	5,527,884	4,207,293	76.11037	1,152,725	20.85291587	154,311	0.027915021	13,555	0.245211368
2000	5,683,047	4,125,263	72.58893	1,356,040	23.8611435	187,490	0.032991105	14,254	0.250816155
2001	5,967,780	4,619,234	77.40289	1,104,439	18.50669763	231,198	0.038741039	12,909	0.216311593
2002	5,271,359	4,329,635	82.13508	868,225	16.47061033	65,143	0.012357914	8,356	0.158516997
2003	6,488,540	5,317,885	81.95811	1,057,804	16.30265052	101,469	0.015638187	11,381	0.175401554
2004	7,129,270	5,566,338	78.07725	1,421,391	19.93739892	127,757	0.017920068	13,784	0.193343779
2005	7,140,596	5,526,773	77.39932	1,466,065	20.53140942	133,730	0.018728129	14,028	0.19645419

* Airline On-Time Performance Data contains only data reported by US certified air carriers that account for at least one percent of domestic scheduled passenger revenues.
* On-time: when a flight arrives less than 15 mins after its published arrival time.

Many interesting observations can be made from the airline on-time performance data. Figure 117 shows average on-time probability of flights that are scheduled to land at ATL for year 1988 and 2005. There is a big drop in late arrivals for year 2005, which can be explained by serious delay propagation problem in later years. Since the airline network becomes more hub-spoke, delays at one point on the NAS propagates throughout the day, resulting in significantly low on-time probability in later times of the day.

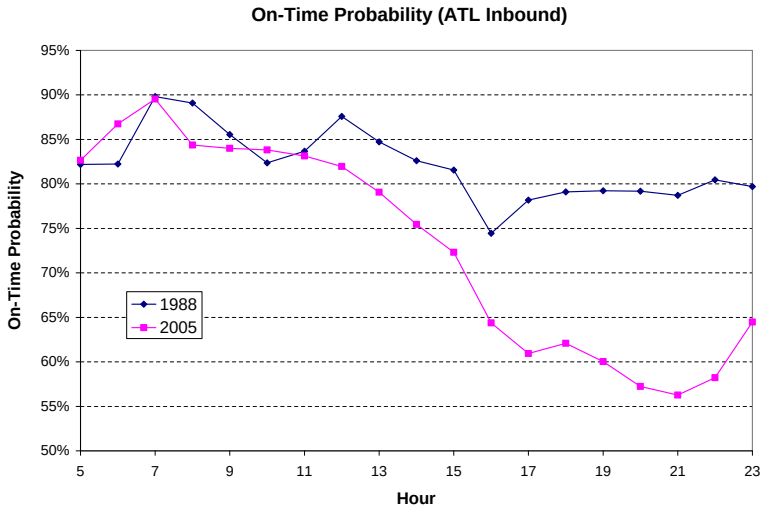


Figure 117: Hourly on-time performance change for ATL inbound flights

Seasonal variations of scheduled block time can also be observed from the data. Figure 118 depicts the average scheduled block time change of operations from ATL to DFW.

Data in big databases tend to include anomalies. Some of the findings that needs attention are reported in this section. It has anomalies in the data as shown in Figure 119. It was found that many of the data has negative air time as shown in Figure 119a. It is suspected that the flights that operates through midnight (day change) is susceptible to this type of error. Also, there are many cells that has zeros or 'nulls'. Figure 119b shows zeros for 'CRS_ARR_TIME' and 'CRS_DEP_TIME' cells. There are also very late or too early flights as can be seen on 'ARR_DELAY' column

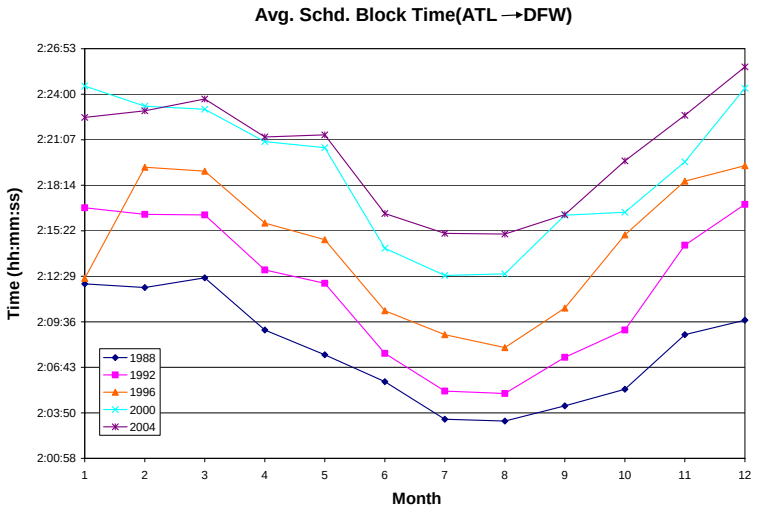


Figure 118: Average monthly scheduled block time change ($ATL \rightarrow DFW$)

in Figure 119c. It shows early arrival of 692 and 939 minutes (top of Figure 119c), which are 11.5 hour and 15.7 hour early. Also, there are arrival delay of 1502, and 1361 minutes (bottom of Figure 119c), which are 25.0 hour and 22.7 hour late. Even if one can claim 25 hour late flight is possible, having 15 hour early flight is hard to accept.

CARRIER	FL_DATE	DAY_OF_WEEK	FLIGHTS	FL_NUM	TAIL_NUM	AIR_TIME	ARR_DELAY	ARR_TIME	CRS_ARR_TIME	DEP_DELAY	DEP_TIME	CRS_DEP_TIME
EV	1/1/2005	6	1	4755	N907EV	-1332	-6	0:50:00	0:56:00	0	23:44:00	23:44:00
EV	1/2/2005	7	1	4755	N854AS	-1331	29	1:25:00	0:56:00	39	0:23:00	23:44:00

(a) Negative air time

CARRIER	FL_DATE	DAY_OF_WEEK	FLIGHTS	FL_NUM	TAIL_NUM	AIR_TIME	ARR_DELAY	ARR_TIME	CRS_ARR_TIME	DEP_DELAY	DEP_TIME	CRS_DEP_TIME
AA	9/1/1996	7	1	1199	N859AA	99	-5	16:29:00	0:00:00	-1	15:29:00	0:00:00
AA	9/2/1996	1	1	1199	N882AA	102	21	16:55:00	0:00:00	15	15:45:00	0:00:00

(b) Lack of data: zeros or nulls

CARRIER	FL_DATE	DAY_OF_WEEK	FLIGHTS	FL_NUM	TAIL_NUM	AIR_TIME	ARR_DELAY	ARR_TIME	CRS_ARR_TIME	DEP_DELAY	DEP_TIME	CRS_DEP_TIME
AS	3/10/2005	4	1	292	N796AS	41	-692	20:28:00	8:00:00	-696	19:24:00	7:00:00
EV	4/22/2005	5	1	4128	N630AS	65	-939	6:20:00	21:59:00	-937	4:00:00	19:37:00

CARRIER	FL_DATE	DAY_OF_WEEK	FLIGHTS	FL_NUM	TAIL_NUM	AIR_TIME	ARR_DELAY	ARR_TIME	CRS_ARR_TIME	DEP_DELAY	DEP_TIME	CRS_DEP_TIME
NW	3/28/2005	1	1	1929	N3322L	79	1502	11:56:00	10:54:00	1517	10:08:00	8:51:00
UA	3/16/2005	3	1	876	N178UA	119	1361	13:34:00	14:53:00	1371	10:16:00	11:25:00

(c) Too early or too late arrival

Figure 119: Sample anomaly in the airline on-time performance data

The airline on-time performance data also shows spikes at every 5 minute period. It is suspected that the people who are in charge of reporting tend to round off times to closest 5 minute intervals. As an example, arrival delay record at ATL is shown

in Figure 120. In the figure, Y axis is in logarithmic scale and the yearly data from 1988 to 2005 were plotted.

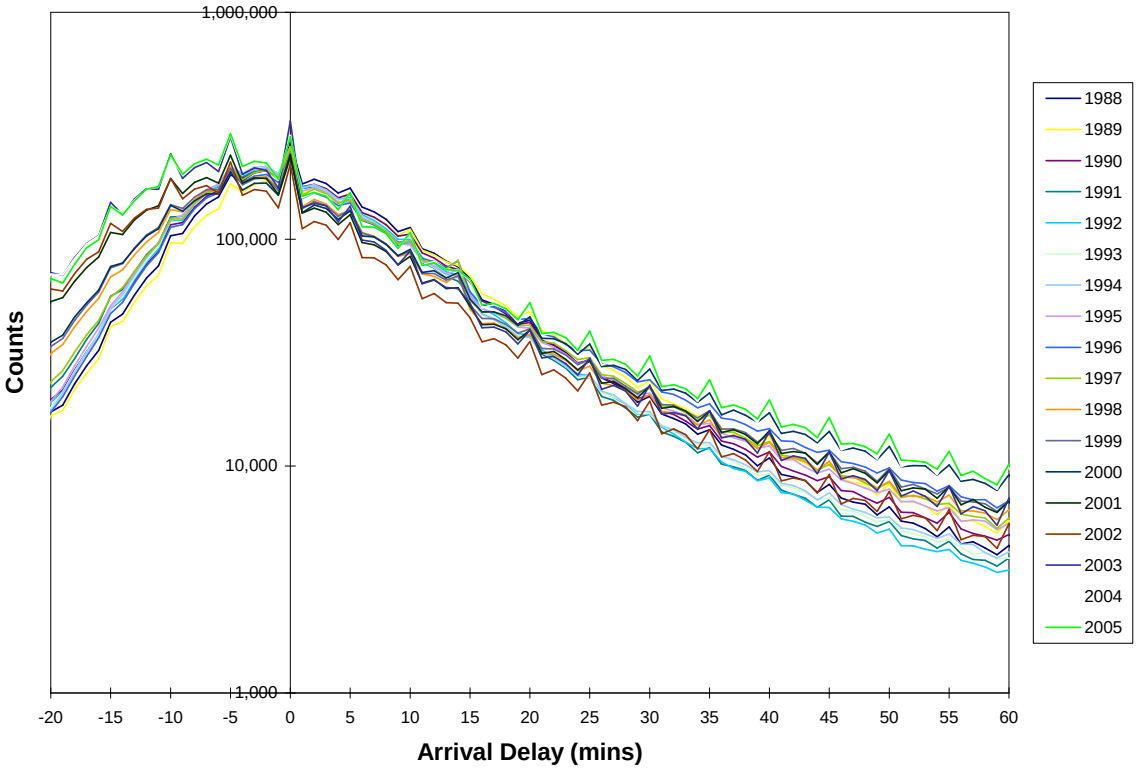


Figure 120: Arrival delay record at ATL

Another irregularity in the database is incorrect N-Numbers. For example, 'N405AA' is registered as balloon on FAA registry[1]. But it made 1,409 scheduled flights in year 2005 with the average speed of around 327 miles/hour according to the airline on-time performance data data. The result of the N-Number inquiry is below.

[1]Available from `http://registry.faa.gov/aircraftinquiry/nnum_inquiry.asp`, last accessed May 2009

FAA Registry N-Number Inquiry Results

N405AA is Assigned

Aircraft Description

Serial Number S55A-252 Type Registration Individual

Manufacturer Name RAVEN Certificate Issue Date 11/08/2006

Model S55A Status Valid

Type Aircraft Balloon Type Engine None

Pending Number Change None Dealer No

Date Change Authorized None Mode S Code 51137465

MFR Year 1975 Fractional Owner NO

Registered Owner

Name THORNHILL ELIZABETH A

Street 67873 30TH AVE

City CATHEDRAL CITY State CALIFORNIA Zip Code 92234-5810

County RIVERSIDE

Country UNITED STATES

Also, there exists non-alphanumeric characters in many of the records, e.g. 'N823?', '?KNO?'. It is suspected that the N-Number of either FAA or BTS's airline on-time performance data is erroneous. An email was sent to Reference Services, National Transportation Library, Bureau of Transportation Statistics, Research and Innovative Technology Administration, U. S. Department of Transportation to enquire about these errors. A partial contents of the reply is in the box below. According to the reply, there is no way to correct these errors.

Thank you for your patience with your inquiry - after some research, the data files are not corrupted but invalid numbers, which sometimes happens for historical data. We downloaded the zip files and found that these invalid numbers only occur for Jan-Feb of 2002, and Jan-December of 2001. Unfortunately there is no way of obtaining this data.

D.2 Enhanced Traffic Management System (ETMS) and Enhanced Traffic Management System Count (ETMSC) Data

The Enhanced Traffic Management System (ETMS) is a traffic prediction and planning tool. Traffic management personnel use ETMS to evaluate traffic demand for all airports and sectors and apply necessary action so that the traffic demand is within capacity. ETMSC provides quick access to traffic counts by airport or by city pair with grouping options such as aircraft type, hour, etc. ETMSC records are created by combining ETMS records with the Aircraft registry data, oceanic data by city pairs provided by MITRE, and OPSNET data. The data covers daily flights from January 2000 and updated daily. The database can be accessed through FAA website (`http://www.apo.data.faa.gov/`). A snap shot of a result from ETMSC is shown in Figure 121.

D.3 Consolidated Operations and Delay Analysis System (CODAS)

The Consolidated Operations and Delay Analysis System (CODAS) has been created to estimate individual flight delay. CODAS integrates databases - mainly the ETMS and the Airline Service Quality Performance (ASQP) System. Additional data from the Official Airline Guide (OAG) and the carriers' Computerized Reservation Systems (CRS) are also used. Aviation System Performance Metrics (ASPM: 2000-Present) is the successor of CODAS (1997-2000).

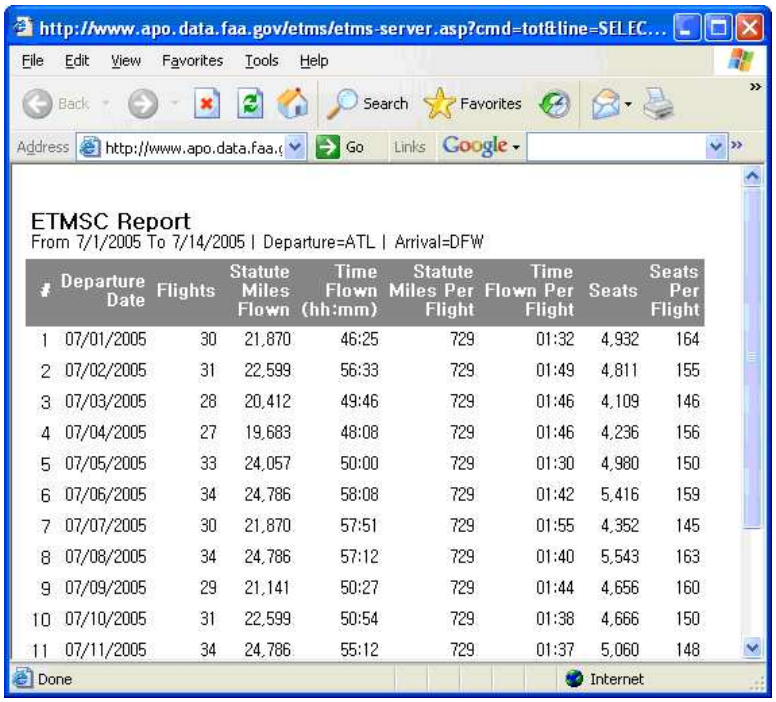


Figure 121: Snapshot - ETMSC

D.4 Aviation System Performance Metrics (ASPM)

The Aviation System Performance Metrics (ASPM) provides information on flight performance and airport efficiency. The data range for airports is from January 2000 to present. Information on 55 airports is available for the entire period and additional 20 airports from October 2004. The data range for arrival/departure rates and runway configuration is from January 1, 2000 and delay information is from January 1, 1998. The database can be accessed through the FAA website (`http://www.apo.data.faa.gov/`). A snap shot of ASPM is shown in Figure 122.

D.5 Airline Service Quality Performance (ASQP)

According to title 14, part 234 of Code of Federal Regulations (CFR), carriers with one percent or more of the total domestic scheduled passenger revenues are required to report flight data that involve any airport in the contiguous US states that account for one percent or more of domestic scheduled enplanements. The

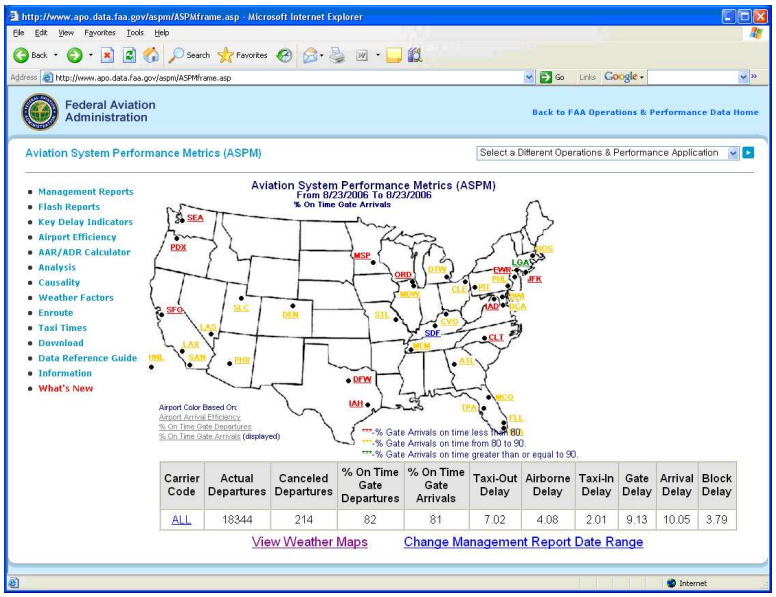
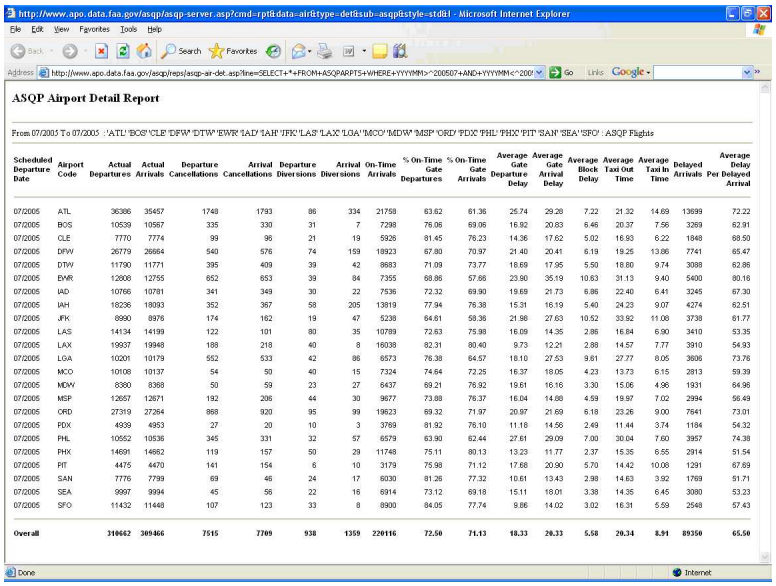


Figure 122: Snapshot - ASPM

data includes gate departure, gate arrival, wheels-off and wheels-on time, causal data for delayed aircrafts, etc. The range of the data covers from June 2003 and monthly update is performed. The database can be accessed through FAA website (`http://www.apo.data.faa.gov/`). A snap shot of a result from the Airline Service Quality Performance (ASQP) is shown in Figure 123.

D.6 Official Airline Guide (OAG)

The Official Airline Guide (OAG) is a private company that maintains and distributes information within the passenger and cargo aviation field. Its airline schedules database contains historical and future flight details for more than 3,500 airports and 1000 airlines. OAG scheduled flight database is used frequently by many research organizations as an independent data source. The website of OAG can be accessed at `http://www.oagflights.com`.



ASQP Airport Detail Report

From 07/2005 To 07/2005 : 'ATL' 'BOS' 'CLE' 'DFW' 'DTW' 'EWR' 'IAD' 'IAH' 'JFK' 'LAS' 'LAX' 'LGA' 'MCO' 'MDW' 'MSP' 'ORD' 'PDX' 'PHL' 'PHX' 'PIT' 'SAN' 'SEA' 'SFO' : ASQP Flights

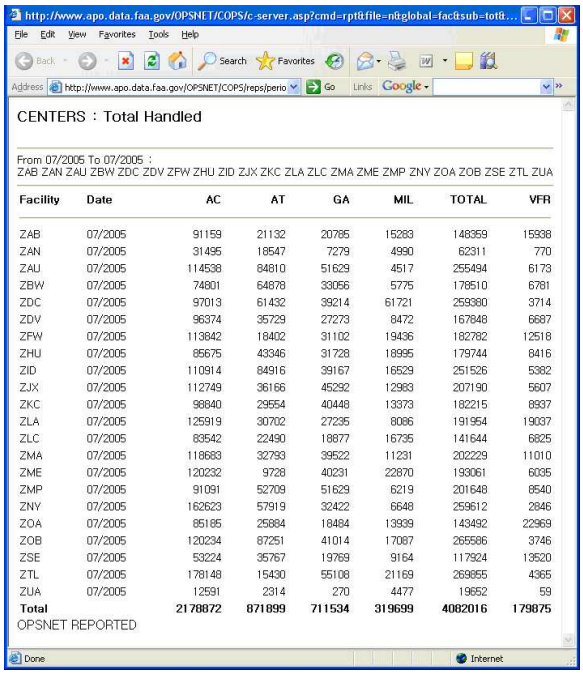
Scheduled Departure Date	Airport Code	Actual Departures	Actual Arrivals	Departure Cancellations	Arrival Cancellations	Departure Diversions	Arrival Diversions	On-Time Arrivals	% On-Time Gate Departures	% On-Time Gate Arrivals	Average Gate Departure Delay	Average Gate Arrival Delay	Average Block Delay	Average Taxi Out Time	Average Taxi In Time	Delayed Arrivals	Average Delay Per Delayed Arrival
07/2005	ATL	36386	35457	1749	1793	86	334	21758	63.62	61.36	25.74	29.28	7.22	21.32	14.69	13699	72.22
07/2005	BOS	10539	10567	335	330	31	7	7298	76.06	69.06	16.92	20.83	6.46	20.37	7.56	3269	62.91
07/2005	CLE	7770	7774	99	98	21	19	5926	81.45	76.23	14.36	17.62	5.02	16.93	6.22	1848	68.50
07/2005	DFW	26779	26664	540	576	74	159	18923	67.80	70.97	21.40	20.41	6.19	19.25	13.86	7741	65.47
07/2005	DTW	11790	11771	395	409	39	42	8683	71.09	73.77	18.69	17.95	5.50	18.80	9.74	3088	62.86
07/2005	EWR	12808	12755	652	653	39	84	7355	68.86	57.66	23.90	35.19	10.63	31.13	9.40	5400	80.16
07/2005	IAD	10766	10781	341	349	30	22	7536	72.32	69.90	19.69	21.73	6.86	22.40	6.41	3245	67.30
07/2005	IAH	18236	18093	352	367	58	205	13819	77.94	76.38	15.31	16.19	5.40	24.23	9.07	4274	62.51
07/2005	JFK	8990	8976	174	162	19	47	5238	64.61	58.36	21.98	27.63	10.52	33.92	11.06	3738	61.77
07/2005	LAS	14134	14199	122	101	80	35	10789	72.63	75.98	16.09	14.35	2.86	16.84	6.90	3410	53.35
07/2005	LAX	19937	19948	188	218	40	8	16038	82.31	80.40	9.73	12.21	2.88	14.57	7.77	3910	54.93
07/2005	LGA	10201	10179	552	533	42	86	6573	76.38	64.57	18.10	27.53	9.61	27.77	8.05	3806	73.76
07/2005	MCO	10108	10137	54	50	40	15	7324	74.64	72.25	16.37	18.05	4.23	13.73	6.15	2813	59.39
07/2005	MDW	8360	8368	50	59	23	27	6437	69.21	76.92	19.61	16.16	3.30	15.06	4.96	1931	64.96
07/2005	MSP	12657	12671	192	206	44	30	9677	73.88	76.37	16.04	14.88	4.59	19.97	7.02	2994	56.49
07/2005	ORD	27319	27264	868	920	95	99	19623	69.32	71.97	20.97	21.69	6.18	23.26	9.00	7641	73.01
07/2005	PDX	4939	4953	27	20	10	3	3769	81.92	76.10	11.18	14.56	2.49	11.44	3.74	1184	54.32
07/2005	PHL	10552	10536	345	331	32	57	6579	63.90	62.44	27.61	29.09	7.00	30.04	7.60	3957	74.38
07/2005	PHX	14691	14662	119	157	50	29	11748	75.11	80.13	13.23	11.77	2.37	15.35	6.55	2914	51.54
07/2005	PIT	4475	4470	141	154	6	10	3179	75.98	71.12	17.68	20.90	5.70	14.42	10.08	1291	67.69
07/2005	SAN	7776	7799	69	46	24	17	6030	81.26	77.32	10.61	13.43	2.98	14.63	3.92	1769	51.71
07/2005	SEA	9997	9994	45	56	22	16	6914	73.12	69.18	15.11	18.01	3.38	14.35	6.45	3080	53.23
07/2005	SFO	11432	11448	107	123	33	8	8900	84.05	77.74	9.86	14.02	3.02	16.31	5.59	2548	57.43
Overall		**310662**	**309466**	**7515**	**7709**	**938**	**1359**	**220116**	**72.50**	**71.13**	**18.33**	**20.33**	**5.58**	**20.34**	**8.91**	**89350**	**65.50**

Figure 123: Snapshot - ASQP

D.7 Operations Network (OPSNET)

The Operations Network (OPSNET) is the official source for air traffic delays and operations. OPSNET includes preliminary airport traffic counts, instrument operations, and instrument approaches, as well as delay and airport traffic information. Daily, monthly, and annual counts are available by facility, state, region, or nationally. The coverage of the data is from 1990 to present and it is updated daily. Historical operations are synchronized with the Air Traffic Activity Data System (ATADS). The database can be accessed through FAA website (`http://www.apo.data.faa.gov/`). A snap shot of a result of 'center' from OPSNET is shown in Figure 124.

The air traffic and delay data provided by BTS is obtained through FAA. The FAA receives daily reports on delays and airport traffic and maintains the information in the OPSNET database. In OPSNET and some other databases, aircraft operations are categorized as 1) Air Carrier, 2) Air Taxi/Commuter, 3) General Aviation, and 4) Military. The definitions and descriptions for these are given in Table 29.



CENTERS : Total Handled

From 07/2005 To 07/2005 :
ZAB ZAN ZAU ZBW ZDC ZDV ZFW ZHU ZID ZJX ZKC ZLA ZLC ZMA ZME ZMP ZNY ZOA ZOB ZSE ZTL ZUA

Facility	Date	AC	AT	GA	MIL	TOTAL	VFR
ZAB	07/2005	91159	21132	20785	15283	148359	15938
ZAN	07/2005	31495	18547	7279	4990	62311	770
ZAU	07/2005	114538	84810	51629	4517	255494	6173
ZBW	07/2005	74801	64878	33056	5775	178510	6781
ZDC	07/2005	97013	61432	39214	61721	259380	3714
ZDV	07/2005	96374	35729	27273	8472	167848	6687
ZFW	07/2005	113842	18402	31102	19436	182782	12518
ZHU	07/2005	85675	43346	31728	18995	179744	8416
ZID	07/2005	110914	84916	39167	16529	251526	5382
ZJX	07/2005	112749	36166	45292	12983	207190	5607
ZKC	07/2005	98840	29554	40448	13373	182215	8937
ZLA	07/2005	125919	30702	27235	8086	191954	19037
ZLC	07/2005	83542	22490	18877	16735	141644	6825
ZMA	07/2005	118683	32793	39522	11231	202229	11010
ZME	07/2005	120232	9728	40231	22870	193061	6035
ZMP	07/2005	91091	52709	51629	6219	201648	8540
ZNY	07/2005	162623	57919	32422	6648	259612	2846
ZOA	07/2005	85185	25884	18484	13939	143492	22969
ZOB	07/2005	120234	87251	41014	17087	265586	3746
ZSE	07/2005	53224	35767	19769	9164	117924	13520
ZTL	07/2005	178148	15430	55108	21169	269855	4365
ZUA	07/2005	12591	2314	270	4477	19652	59
Total		**2178872**	**871899**	**711534**	**319699**	**4082016**	**179875**

OPSNET REPORTED

Figure 124: Snapshot -OPSNET - center

Table 29: Categories of aircraft operation. [Source: Federal Aviation Administration (2009*a*)]

Category	**Definition**
Air Carrier (AC)	An aircraft with seating capacity of more than 60 seats or a maximum payload capacity of more than 18,000 pounds carrying passengers or cargo for hire or compensation. This includes US and foreign flagged carriers.
Air Taxi (AT)/ Commuter (Comm.)	Aircraft designed to have a maximum seating capacity of 60 seats or less or a maximum payload capacity of 18,000 pounds or less carrying passengers or cargo for hire or compensation.
General Aviation (GA)	Takeoffs and landings of all civil aircraft, except those classified as air carriers or air taxis.
Military (Mil)	All classes of military takeoffs and landings at FAA and FAA-contracted facilities.

D.8 Air Traffic Activity Data System (ATADS)

The Air Traffic Activity Data System (ATADS) provides air traffic operations counts for towered airports, TRACONS, centers as well as instrument and approach counts. Data can be obtained for daily, monthly, and annual counts categorized by facility, state, region, or nationally. The official source of the data is the OPSNET. The database can be accessed through FAA website (`http://www.apo.data.faa.gov/`).

D.9 Terminal Area Forecast (TAF)

The Terminal Area Forecast(TAF) is the official enplanements and operations forecast at FAA facilities. Facilities in the forecast are (1) FAA towered airports, (2) Federally contracted towered airports, (3) Nonfederal towered airports, and (4) Non-towered airports. The forecasts are made for four main categories, which are (1) Large air carriers, (2) Air taxi/commuters, (3) General aviation, and (4) Military.

Since significant changes are being made in the aviation industries, the forecasts are being revised from time to time. Therefore the forecasts need to be accessed when the update is current. The database can be accessed through FAA website (`http://www.apo.data.faa.gov/`). A snap shot of a result from TAF is shown in Figure 125.

D.10 Airport delay information services

FAA provides many useful information services under Air Traffic Control System Command Center (ATCSCC). Real time flight delay information is provided from its' main page, which can be accessed through FAA website (`http://www.fly.faa.gov/flyfaa/usmap.jsp`). Also, real time arrival demand information for selected airports can be obtained through AIRPORT ARRIVAL DEMAND CHART (AADC), which can be accessed at (`http://www.fly.faa.gov/Products/AADC/aadc.html`). As in Figure 126, users can select number of airports from the combo box menu.

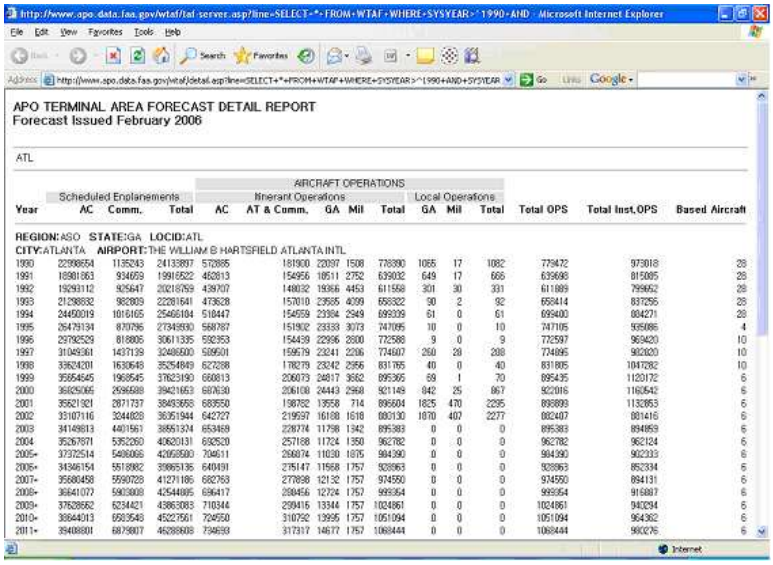


Figure 125: Snapshot - Terminal Area Forecast

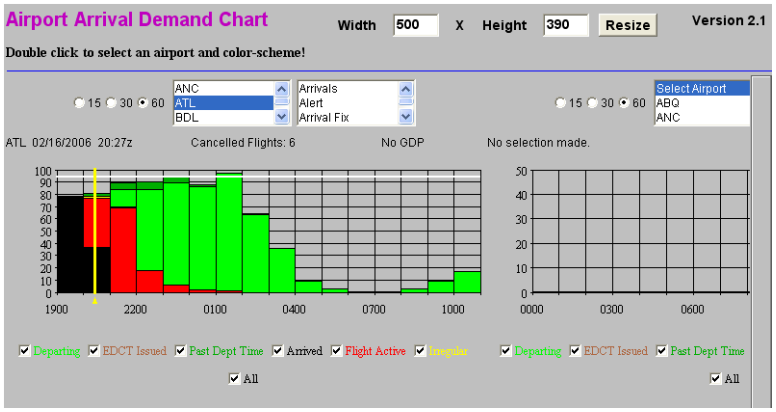


Figure 126: Screen capture from AADC

Bureau of Transportation Statistics(BTS) has a web based service on airline on-time statistics and delay causes. It can be accessed at `http://www.transtats.bts.gov/OT_Delay/OT_DelayCause1.asp` It also provides detailed data on the same topic, which can be interactively generated from the database depending on the users choice. Figure 127, shows a screen capture from the web.

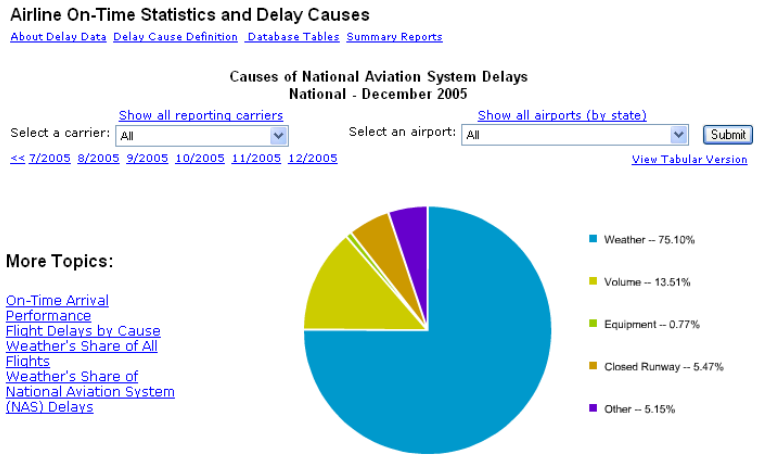


Figure 127: Airline on-time statistics and delay causes. [Source: BTS website]

APPENDIX E

THE OPERATIONAL EVOLUTION PARTNERSHIP

The multi-agency Joint Planning and Development Office (JPDO) is developing the vision for the Next Generation Air Transportation System (NextGen). Each member of the JPDO will work on their individual plans. The Operational Evolution Partnership (OEP) is the FAA's plan for implementing the NextGen (Federal Aviation Administration, 2007). The OEP is an expansion of the original OEP (the Operational Evolution Plan), which started in year 2001. OEP version 1 describes the implementation plan, which are divided into three domains:

- Airport development: Capacity enhancements and reducing delays at airports
- Aircraft and operator requirements: Development of standards for the new capabilities required by NextGen
- Air traffic operations: Developing air traffic control capabilities

The 35 airports included in the OEP accounts for 75 percent of all passenger enplanements (Federal Aviation Administration, 2007). They are the most heavily traveled airports located in the most densely populated areas. It was determined that most of the current delay on the NAS is traced to inadequate throughput (i.e. departure and arrival rates) at some of the busiest airports. Figure 128 shows the locations of OEP35 airports. Enplanements at OEP35 airports from year 2001 is shown in Table 30

The most significant capacity increases can come from the construction of new runways and runway extensions. Since year 2000, 13 new runways have been built at the OEP35 airports, which increased the capacity to accommodate 1.6 million more

Table 30: Enplanements at OEP35 airports. [Data source: FAA Passenger and All-Cargo Statistics]

Airport	2001	2002	2003	2004	2005
ATL	37,181,068	37,720,556	38,893,670	41,123,857	42,402,653
BWI	10,098,665	9,367,499	9,768,040	10,103,563	9,829,432
BOS	11,739,553	11,077,238	11,087,799	12,758,020	13,214,923
CLT	11,548,952	11,743,157	11,465,366	12,499,476	14,009,608
MDW	7,112,784	7,878,438	8,687,215	9,238,592	8,383,698
ORD	31,529,561	31,706,328	32,920,387	36,100,147	36,720,005
CVG	8,586,907	10,316,170	10,449,930	10,864,547	11,277,068
CLE	5,633,495	5,146,975	5,012,446	5,389,196	5,529,629
DFW	25,610,562	24,761,105	24,976,881	28,063,035	28,079,147
DEN	17,178,872	16,943,564	17,969,754	20,407,002	20,799,886
DTW	15,819,584	15,525,413	15,754,017	17,046,176	17,580,363
FLL	8,015,055	8,266,788	8,682,781	10,040,598	10,729,468
IAH	16,173,551	15,865,479	16,134,684	17,322,065	19,032,196
PIT	9,939,223	8,975,111	7,113,460	6,606,117	5,198,442
HNL	9,810,860	9,406,467	9,044,409	9,579,076	9,784,404
STL	13,264,751	12,474,566	9,922,456	6,377,628	6,844,769
LAS	16,633,435	16,600,807	17,097,738	19,943,025	21,402,688
LAX	29,365,436	26,911,570	26,239,584	28,925,341	29,372,253
MEM	5,560,524	5,231,998	5,411,496	5,295,062	5,630,305
MIA	14,941,663	14,020,686	14,198,321	14,515,591	15,092,763
MSP	15,852,433	15,544,039	16,022,988	17,482,627	17,971,771
JFK	14,553,815	14,552,411	15,676,352	18,586,863	20,199,967
LGA	11,352,248	11,076,032	11,367,309	12,312,561	12,983,410
EWR	15,497,560	14,553,843	14,628,708	15,827,675	16,433,587
MCO	13,622,397	12,921,480	13,375,162	15,270,347	16,592,133
PHL	11,736,129	11,954,469	11,870,928	13,824,332	15,376,569
PHX	17,478,622	17,271,519	18,252,853	19,336,099	20,315,542
PDX	6,168,103	5,978,025	6,059,860	6,379,884	6,798,976
DCA	6,267,395	6,172,065	6,813,148	7,661,532	8,623,907
SLC	8,951,776	8,997,942	8,958,003	8,884,880	10,601,918
SAN	7,506,320	7,392,389	7,565,196	8,135,832	8,628,648
SFO	16,475,611	14,736,137	14,079,173	15,605,822	16,070,127
SEA	13,184,630	12,969,024	13,109,153	14,092,285	14,359,487
TPA	7,901,725	7,726,576	7,672,533	8,436,025	9,297,643
IAD	8,484,112	7,848,911	8,050,506	10,961,614	13,032,502
Total	480,777,377	469,634,777	474,332,306	514,996,492	538,199,887

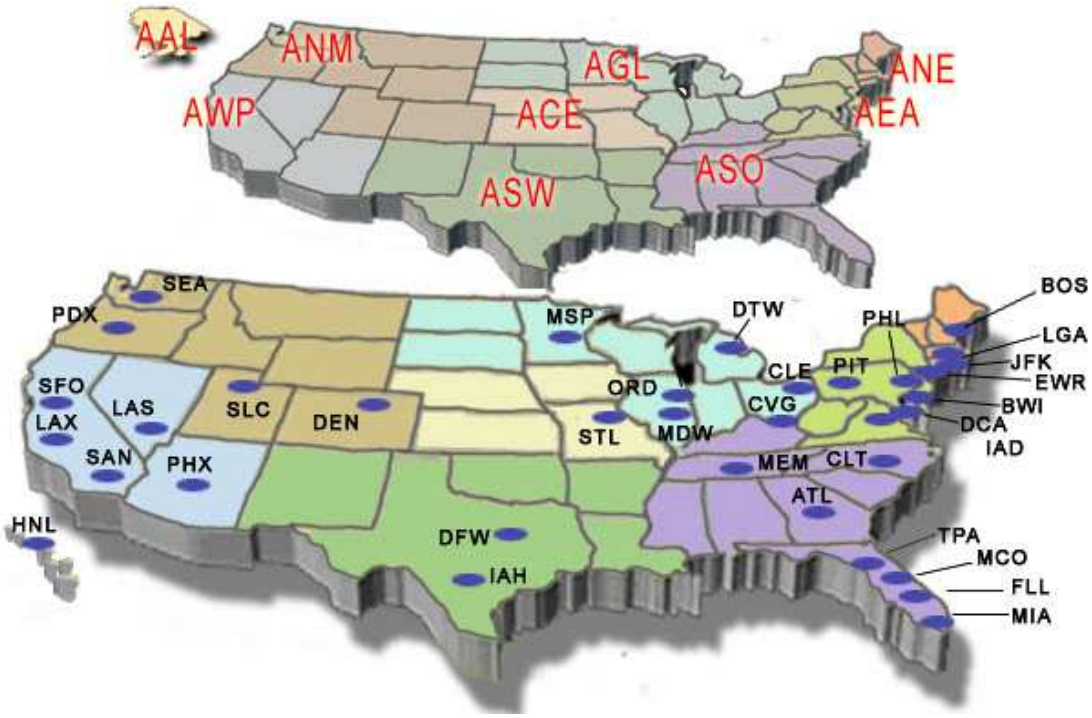


Figure 128: OEP 35 benchmark airports. [Source: FAA website]

annual operations. There are ten airfield projects at eight of the OEP35 airports at the moment. These ten core OEP airports projects will be commissioned through 2011 and it is expected that the potential to accommodate operations at these airports will be increased by about 400,000 annual operations. Table 31 shows new runway projects at 8 OEP35 airports. More information about OEP35 airports are given in Table 32.

Table 31: Runway construction at eight of OEP35 airports. [Source: Federal Aviation Administration (2007)]

Location ID	**City**	**Project**	**To Open**	**Operational Benefit**	
				Increase in Annual Ops	**Delay Reduction (per op., (min.))**
LAX	Los Angeles	Relocated Runway New Center Taxiway	April 2 2007 Jun 24 2008		
SEA	Seattle	New 8,500' Runway	Nov 20 2008	175,000	3.4
IAD	Washington	New 9,400' Runway	Nov 20 2008	70,000	2.5
ORD	Chicago	New 7,500' Runway 2,856' Runway Extension Relocated 10,800' RW	Nov 20 2008 Nov 20 2008 2011	60,000	0.7
DFW	Dallas-Ft. Worth	End Around Taxiway	Dec 14 2008		
PHL	Philadelphia	1,040' Runway Extension	Mar-09		1.4
BOS	Boston	9,300' Centerfield Taxiway	Nov-09		Ground delays by up to 22%
CLT	Charlotte	New 9,000' Runway	Feb 11 2010	80,000	1.5

Table 32: Activities at OEP35 airports in the U.S. in 2002. [Source: Bhadra and Texter (2004)]

Airport code	Primary and secondary air carriers	City	State	Hub type	Number of annual scheduled passengers	Percentage of total passengers	Number of annual scheduled A/C ops	Percentage of total A/C ops
ATL	Delta, Airtran	Atlanta	GA	L	36,321,239	5.79	642,727	4.87
ORD	United, American	Chicago	IL	L	31,026,878	4.94	612,553	4.64
LAX	United, American, Southwest	Los Angeles	CA	L	26,323,259	4.19	447,170	3.39
DFW	American, Delta	Dallas-Ft Worth	TX	L	24,148,619	3.85	493,772	3.74
PHX	America West, Southwest	Phoenix	AZ	L	16,930,419	2.70	370,247	2.80
DEN	United, Frontier	Denver	CO	L	16,544,458	2.64	330,825	2.50
LAS	Southwest, America West, United	Las Vegas	NV	L	16,540,417	2.64	317,700	2.41
IAH	Continental, American	Houston	TX	L	15,889,349	2.53	299,903	2.27
MSP	Northwest, American	Minneapolis	MN	L	15,553,423	2.48	326,974	2.48
DTW	Northwest	Detroit	MI	L	15,124,490	2.41	337,816	2.56
JFK	American, Jet Blue, Delta, United	New York	NY	L	15,050,456	2.40	245,475	1.86
SFO	United, American	San Francisco	CA	L	14,856,842	2.37	260,501	1.97
EWR	Continental, American	Newark	NJ	L	14,073,453	2.24	282,849	2.14
MIA	American, Continental	Miami	FL	L	13,889,275	2.21	304,863	2.31
MCO	Delta, Southwest, American	Orlando	FL	L	12,631,347	2.01	201,203	1.52
SEA	Alaska, United	Seattle	WA	L	12,570,572	2.00	217,352	1.65
STL	American, Southwest	St Louis	MO	L	12,412,120	1.98	295,148	2.23
PHL	US Airways, American	Philadelphia	PA	L	11,631,738	1.85	267,402	2.02
CLT	US Airways	Charlotte	NC	L	11,589,824	1.85	239,173	1.81
BOS	American, US Airways, Delta	Boston	MA	L	10,665,476	1.70	207,138	1.57
LGA	US Airways, American, Delta	New York	NY	L	10,416,041	1.66	207,915	1.57
CVG	Delta	Covington-Cincinnati, OH	KY	L	9,879,246	1.57	150,943	1.14

Continued on next page

Table 32 – continued from previous page

Airport code	Primary and secondary air carriers	City	State	Hub type	Number of annual scheduled passengers	Percentage of total passengers	Number of annual scheduled A/C ops	Percentage of total A/C ops
BWI	Southwest, US Airways	Baltimore	MD	L	9,489,296	1.51	210,349	1.59
HNL	Aloha, Hawaiian, ATA	Honolulu	HI	L	9,266,615	1.48	174,544	1.32
PIT	US Airways, Delta	Pittsburgh	PA	L	9,212,335	1.47	188,154	1.42
SLC	Delta, Southwest	Salt Lake City	UT	L	8,908,798	1.42	151,121	1.14
FLL	Continental, Delta, Southwest, American, US Airways, Spirit, Jet Blue	Ft Lauderdale	FL	L	7,983,704	1.27	147,874	1.12
IAD	United, Delta, American	Washington	DC	L	7,936,618	1.26	130,351	0.99
MDW	American, Southwest	Chicago	IL	L	7,573,932	1.21	161,468	1.22
TPA	Southwest, Delta, US Airways, Continental, Air Tran, American	Tampa	FL	L	7,544,284	1.20	145,968	1.11
SAN	Southwest, American, United	San Diego	CA	L	7,224,573	1.15	143,298	1.08
PDX	Alaska, United, Southwest	Portland	OR	M	5,970,960	0.95	122,407	0.93
DCA	US Airways, Delta	Washington	DC	M	5,311,436	0.85	121,456	0.92
CLE	Continental, American	Cleveland	OH	M	5,007,767	0.80	86,572	0.66
MEM	Northwest	Memphis	TN	M	4,784,135	0.76	237,385	1.80

Key: A/C = aircraft; L = large hubs; M = medium hubs.

Hub types in the table is defined by the U.S. Department of Transportation, Federal Aviation Administration. There are four categories depending on the total enplanements, which are 1) large ($> 1\%$ of total enplanements), 2) medium ($0.25\% - 0.999\%$ of total enplanements), 3) small hubs ($0.05\% - 0.249\%$ of total enplanements), and 4) nonhub ($< 0.05\%$ of total enplanements).

Bibliography

Adler, N. (2005), 'Hub-spoke network choice under competition with an application to western europe', *Transportation Science* **39**, 58–72.

Adler, N. and Berechman, J. (2001), 'Evaluating optimal multi-hub networks in a deregulated aviation market with an application to western Europe', *Transportation Research Part A* **35**, 373–390.

Airbus S.A.S. (2007), 'Global market forecast 2007 – 2026'. Available from `http://www.airbus.com/en/corporate/gmf/`, last accessed May 2009.

Albert, R. and Barabsi, A.-L. (2002), 'Statistical mechanics of complex networks', *Review of Modern Physics* **74**, 47–97.

Albert, R., Jeong, H. and Barabási, A.-L. (1999), 'Internet: Diameter of the world-wide web', *Nature* **401**(6749), 130–131.

Albert, R., Jeong, H. and Barabási, A.-L. (2000), 'Error and attack tolerance of complex networks', *Nature* **406**(6794), 378–382.

Alderighi, M., Cento, A., Nijkamp, P. and Rietveld, P. (2005), 'Network competition — the coexistence of hub-and-spoke and point-to-point systems', *Journal of Air Transport Management* **11**, 328–334.

Alderighi, M., Cento, A., Nijkamp, P. and Rietveld, P. (2007), 'Assessment of new hub-and-spoke and point-to-point airline network configurations', *Transport Reviews* **27**(5), 529–549.

Alexandrov, N. (2004), Transportation network topologies, Technical Report TM-2004-213259, NASA Langley and National Institute of Aerospace.

Baik, H. and Trani, A. A. (2005), 'A Transportation Systems Analysis Model (TSAM) to study the impact of the Small Aircraft Transportation System (SATS)'. Available from `http://www.atsl.cee.vt.edu/tsam.htm`, last accessed May 2009.

Bania, N., Bauer, P. W. and Zlatoper, T. J. (1998), 'US air passenger service: a taxonomy of route networks, hub locations, and competition', *Logistics and Transportation Review* **34**(1), 53–74.

Barabási, A.-L. (2004), *Linked: How Everything is Connected to Everything Else*, Basic Books. ISBN: 0-452-28439-2.

Barabási, A.-L. and Albert, R. (1999), 'Emergence of scaling in random networks', *Science* **Vol 286, Issue 5439**, 509–512.

Barabási, A.-L., Albert, R. and Jeong, H. (1999), 'Mean-field theory for scale-free random networks', *Physica A* **272**, 173–187.

Barabási, A.-L. and Bonabeau, E. (2003), 'Scale-free networks', *Scientific American* **288**, 60–69.

Barabási, A.-L. and Oltvai, Z. N. (2004), 'Network biology: Understanding the cells's functional organization', *Nature Reviews Genetics* **5**, 101–113.

Barla, P. and Constantatos, C. (2000), 'Airline network structure under demand uncertainty', *Transportation Research Part E* **36**, 173–180.

Bhadra, D. (2003), 'Demand for air travel in the United States: Bottom-up econometric estimation and implications for forecasts by origin and destination pairs', *Journal of Air Transportation* **8**(2), 19–56.

Bhadra, D., Gentry, J., Hogan, B. and Wells, M. (2003), 'Building the timetable from bottom-up demand: A micro-econometric approach', MITRE technical paper.

Bhadra, D. and Texter, P. (2004), 'Airline networks: An econometric framework to analyze domestic U.S. air travel', *Journal of Transportation and Statistics* **7**(1), 87–102.

Bohn, R. E. (1994), 'Measuring and managing technological knowledge', *Sloan Management Review* **36**(1), 61–73.

Bonnefoy, P. A. and Hansman, Jr, R. J. (2007), Scalability and evolutionary dynamics of air transportation networks in the United States, number AIAA 2007-7773, AIAA Aviation Technology, Integrations, and Operations Conference.

Borener, S., Carr, G., Ballard, D. and Hasan, S. (2006), 'Can NGATS meet the demands of the future?', *Journal of Air Traffic Control* **48, Issue 1**, 34–38.

Borgatti, S. P. (2005), 'Centrality and network flow', *Social Networks* **27**, 55–71.

Braha, D. and Bar-Yam, Y. (2004), 'Topology of large-scale engineering problem-solving networks', *Physical Review E* **69**, 016113.

Brueckner, J. K. (2004), 'Network structure and airline scheduling', *The Journal of Industrial Economics* **52**, 291–312.

Brueckner, J. K., Dyer, N. J. and Spiller, P. T. (1992), 'Fare determination in arline hub-and-spoke networks', *The RAND Journal of Economics* **23**(3), 309–333.

Brueckner, J. K. and Zhang, Y. (2001), 'A model of scheduling in airline networks', *Journal of Transport Economics and Policy* **35**, 195–222.

Bureau of Transportation Statistics (1998), 'Office of inspector general audit report: Passenger origin-destination data submitted by air carriers', Report Number AV-1998-086.

Bureau of Transportation Statistics (2008), 'Airline origin and destination survey (DB1B)'. [online database], Available from `http://www.transtats.bts.gov/`, last accessed May 2009.

Bureau of Transportation Statistics (2009*a*), 'Air carrier statistics (form 41 traffic) - U.S. carriers'. [online database], Available from `http://www.transtats.bts.gov/`, last accessed May 2009.

Bureau of Transportation Statistics (2009*b*), 'Airline on-time performance data'. [online database], Available from `http://www.transtats.bts.gov/`, last accessed May 2009.

Burghouwt, G. and Hakfoort, J. (2001), 'The evolution of the european aviation network', *Journal of Air Transport Management* **7**, 311–318.

Burghouwt, G., Hakfoort, J. and van Eck, J. R. (2003), 'The spatial configuration of airline networks in europe', *Journal of Air Transport Management* **9**, 309–323.

Button, K. (2002), 'Debunking some common myths about airport hubs', *Journal of Air Transport Management* **8**, 177–188.

Caldarelli, G. (2007), *Scale-Free Networks*, Oxford University Press. ISBN13: 978-0-19-921151-7.

Caldarelli, G., Capocci, A., Rios, P. D. L. and Muñoz, M. (2002), 'Scale-free networks without growth or preferential attachment: Good get richer', cond-mat/0207366v2.

Caves, D. W., Christensen, L. R. and Tretheway, M. W. (1984), 'Economies of density versus economies of scale: Why trunk and local service airline costs differ', *The RAND Journal of Economics* **15**(4), 471–489.

Clarke, J.-P., Melconian, T., Bly, E. and Rabbani, F. (2007), 'MEANS MIT Extensible Air Network Simulation', *Simulation* **83**(5), 385–399.

Cohen, R. and Havlin, S. (2003), 'Scale-free networks are ultrasmall', *Physical Review Letters* **90**(5), 058701.

Couluris, G. J., Hunter, C. G., Blake, M. and Roth, K. (2003), National airspace system simulation capturing the interactions of air traffic management and flight trajectories, AIAA Modeling and Simulation Technologies Conference and Exhibit. AIAA 2003-5597.

DeLaurentis, D., Han, E.-P. and Kotegawa, T. (2008), 'Network-theoretic approach for analyzing connectivity in air transportation networks', *Journal of Aircraft* **45**(5), 1669–1679.

Department of Transportation (1999), 'The 1995 American Travel Survey', Bureau of Transportation Statistics. Microdata CD-ROM.

Department of Transportation (2004), 'Short- and long-term efforts to mitigate flight delays and congestion', Report Number: CR-2004-066.

Department of Transportation (2005), 'Aviation data modernization; proposed rules', Federal Register Notice, Volume 70, Number 32.

Department of Transportation (2009), 'Air travel consumer report'. Available from `http://airconsumer.ost.dot.gov/reports/index.htm`, last accessed May 2009.

Department of Transportation, Federal Aviation Administration and The MITRE Corporation (2004), Capacity needs in the National Airspace System: An analysis of airport and metropolitan area demand and operational capacity in the future, Technical report. Available from `http://www.dot.gov/affairs/links062404.htm`, last accessed May 2009.

Desaulniers, G., Desrosiers, J., Dumas, Y., Solomon, M. M. and françois Soumis (1997), 'Daily aircraft routing and scheduling', *Management Sciences* **43**, 841–855.

Detroit Metropolitan Wayne County Airport and University of Michigan (2006), Detroit Metropolitan Wayne County Airport 2006 economic impact study, Technical report.

Dobson, G. and Lederer, P. J. (1993), 'Airline scheduling and routing in a hub-and-spoke system', *Transportation Science* **27**, 281–297.

Donohue, G. L. and Zellweger, A. G., eds (2001), *Air Transportation Systems Engineering*, AIAA. ISBN-13: 978-1563474743.

Dorogovtsev, S. N. and Mendes, J. F. F. (2002), 'Evolution of networks', *Advances in Physics* **51**(4), 1079–1187.

Dorogovtsev, S. N., Mendes, J. F. F. and Samukhin, A. N. (2000), 'Structure of growing networks with preferential linking', *Physical Review Letters* **85**(21), 4633–4636.

Erdös, P. and Rényi, A. (1959), 'On random graphs', *Publicationes Mathematicae* **6**, 290–297.

Erdös, P. and Rényi, A. (1960), 'On the evolution of random graphs', *Hungarian Academy of Sciences* **5**, 17–61.

Faloutsos, M., Faloutsos, P. and Faloutsos, C. (1999), 'On power-law relationships of the internet topology', *Computer Communication Review* **29**(4), 251–262.

Federal Aviation Administration (2001), 'Investment analysis benefit guidelines: Quantifying flight efficiency benefits'. Version 3.0.

Federal Aviation Administration (2004), Airport capacity benchmark report, Technical report. Available from `http://www.faa.gov/about/office_org/headquarters_offices/ato/publications/bench/`, last accessed May 2009.

Federal Aviation Administration (2007), 'Operational Evolution Partnership (OEP)', Version 1.

Federal Aviation Administration (2008*a*), 'FAA Aerospace Forecasts, Fiscal Years 2008-2025'. Available from `http://www.faa.gov/data_research/aviation/aerospace_forecasts/`, last accessed May 2009.

Federal Aviation Administration (2008*b*), 'Terminal area forecast summary, fiscal years 2007-2025'. Available from `http://www.faa.gov/data_research/aviation/taf_reports/`, last accessed May 2009.

Federal Aviation Administration (2009*a*), 'Operations Network (OPSNET)'. [online database], Available from `http://aspm.faa.gov/opsnet/sys/main.asp`, last accessed May 2009.

Federal Aviation Administration (2009*b*), 'Terminal Area Forecast (TAF)'. [online database], Available from `http://aspm.faa.gov/main/taf.asp`, last accessed May 2009.

Flyvbjerg, B., Holm, M. K. S. and Buhl, S. L. (2006), 'Inaccuracy in traffic forecasts', *Transport Reviews* **26**(1), 1–24.

Freeman, L. (1978), 'Centrality in social networks', *Social Networks* **1**, 215–239.

Garber, N. J. and Hoel, L. A. (2002), *Traffic & Highway Engineering*, 3rd edn, Brooks/Cole.

González, M. C. (2006), Contact Networks of Mobile Agents and Spreading Dynamics, PhD thesis, University of Stuttgart, Germany.

Hansen, M. and Weidner, T. (1995), 'Multiple airport systems in the United States: current status and future prospects', *Transportation Research Record* (1506), 8–17.

Hsiao, C.-Y. and Hansen, M. (2005), 'Air transportation network flows: Equilibrium model', *Transportation Research Record* **1915**, 12–19.

Hsu, C.-I. and Wen, Y.-H. (2000), 'Application of grey theory and multiobjective programming towards airline network design', *European Journal of Operational Research* **127**, 44–68.

Huang, A. S., Schleicher, D. and Hunter, G. (2004), Future flight demand generation tool, number AIAA-2004-6400, AIAA 4th Aviation Technology, Integration and Operations Forum.

Huang, A. and Schleicher, D. (2007), A flexible demand generation application for NextGen concept evaluation, number AIAA-2007-6876.

Huberman, B. A. and Adamic, L. A. (1999), 'Internet: Growth dynamics of the World-Wide Web', *Nature* **401**(6749), 131.

Hunter, G., Ramamoorthy, K., Cobb, P., Huang, A., Blake, M. and Klein, A. (2005), Evaluation of future National Airspace System architectures, number AIAA-2005-6492, AIAA Modeling and Simulation Technologies Conference and Exhibit.

Jeong, H., Mason, S., Barabási, A.-L. and Oltvai, Z. N. (2001), 'Lethality and centrality in protein networks', *Nature* **411**, 41–42.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. and Barabási, A.-L. (2000), 'The large-scale organization of metabolic networks', *Nature* **407**, 651–654.

Joint Planning and Development Office (2007), Concept of operations for the next generation air transportation system, Technical Report Version 1.2.

Kotegawa, T. and DeLaurentis, D. (2007), Evolution of service provider behaviors via network-based analysis, number AIAA 2007-7755, AIAA Aviation Technology, Integration, and Operations Conference.

Krapivsky, P. L., Redner, S. and Leyvraz, F. (2000), 'Connectivity of growing random networks', *Physical Review Letters* **85**(21), 4629–4632.

Lederer, P. J. and Nambimadom, R. S. (1998), 'Airline network design', *Operations Research* **46**(6), 785–804.

Lewe, J.-H. (2005), An Integrated Decision-Making Framework for Transportation Architectures: Application to Aviation Systems Design, PhD thesis, Georgia Institute of Technology.

Lewe, J.-H. (2008), 'Active control algorithms of the air transportation network', presentation in 16^{th} Annual External Advisory Board Review, Aerospace Systems Design Laboratory, Georgia Institute of Technology.

Lewe, J.-H., DeLaurentis, D. A., Mavris, D. N. and Schrage, D. P. (2006), 'Entity-centric abstraction and modelling framework for transportation architectures', *Journal of Air Transportation* **11**(3), 3–33.

Li, L., Alderson, D., Tanaka, R., Doyle, J. C. and Willinger, W. (2005), Towards a theory of scale-free graphs: Definition, properties, and implications (extended version), Technical Report CIT-CDS-04-006, California Institute of Technology.

Lim, S. (2008), An Integrative Assessment of the Commercial Air Transportation System via Adaptive Agents, PhD thesis, Georgia Institute of Technology.

Lohatepanont, M. and Barnhart, C. (2004), 'Airline schedule planning: Integrated models and algorithms for schedule design and fleet assignment', *Transportation Science* **38**(1), 19–32.

Long, D., Lee, D., Gaier, E., Johnson, J. and Kostiuk, P. (1999), A method for forecasting commercial air traffic schedule in the future, Technical Report NASA/CR-1999-208987, NASA Langley Research Center.

Long, D., Lee, D., Johnson, J., Gaier, E. and Kostiuk, P. (1999), Modeling air traffic management technologies with a queuing network model of the National Airspace System, Technical Report NASA/CR-1999-208988, NASA Langley Research Center.

Mane, M., Crossley, W. A. and Nusawardhana, A. (2007), 'System-of-systems inspired aircraft sizing and airline resource allocation via decomposition', *Journal of Aircraft* **4**, 1222–1235.

Marcotte, P. and Nguyen, S., eds (1998), *Equilibrium and advanced transportation modelling*, Kluwer Academic Publishers. ISBN: 0-7923-8162-9.

Mendelson, H. (2003), 'Metcalfe's law and network effects', *Graduate School of Business, Stanford University* .

Merlis, E. A. (2001), 'Before the senate aviation subcommittee hearing on aviation delay prevention', Senate Aviation Subcommittee hearing.

Morrell, P. (2005), 'Airlines within airlines: An analysis of US network airline responses to low cost carriers', *Journal of Air Transport Management* **11**(5), 303–312.

National Aeronautics and Space Administration (2003), 'FACET: Future ATM Concepts Evaluation Tool fact sheet', Advanced Air Transportation Technologies (AATT) Project. Available from `http://as.nasa.gov/aatt/FACET_Oct2003.pdf`, last accessed May 2009.

National Aeronautics and Space Administration (2004), 'The Virtual Airspace Modeling and Simulation (VAMS) fact sheet'. Available from `http://vams.arc.nasa.gov/`, last accessed May 2009.

Newman, M. E. J. (2003), 'The structure and function of complex networks', *SIAM Review* **45**(2), 167–256.

Newman, M. E. J. (2004), 'Power laws, Pareto distributions and Zipf's law', *Contemporary Physics* **46**, 323351.

Niedringhaus, W. P. (2004), 'The Jet:Wise model of National Airspace System evolution', *Simulation* **80**(1), 45–58.

O'Connor, C. J. and Rutishauser, D. K. (2001), Enhanced airport capacity through safe, dynamic reductions in aircraft separation: NASA's Aircraft VOrtex Spacing System (AVOSS), Technical Report NASA/TM-2001-211052, NASA Langley Research Center.

O'Kelly, M. E. and Miller, H. J. (1994), 'The hub network design problem', *Journal of Transport Geography* **2**(1), 31–40.

Onody, R. N. and de Castro, P. A. (2004), 'Complex network study of Brazilian soccer players', *Physical Review E* **70**(3), 037103.

Oum, T., Zhang, A. and Zhang, Y. (1995), 'Airline network rivalry', *Canadian Journal of Economics* **28**(4a), 836–857.

Pastor-Satorras, R. and Vespignani, A. (2001), 'Epidemic dynamics and endemic states in complex networks', *Physical Review E* **63**, 066117.

Poole, Jr, R. W. and Butler, V. (1999), 'Airline deregulation: The unfinished revolution', *Regulation* **22**(1), 44–51. Available from `http://www.cato.org/pubs/regulation/regv22n1/airline.pdf`, last accessed May 2009.

Potts, R. B. and Oliver, R. M. (1972), *Flows in Transportation Networks*, Academic Press. ISBN: 0125636504.

Price, D. (1976), 'A general theory of bibliometric and other cumulative advantage processes', *Journal of the American Society for Information Science* **27 (5-6)**, 292–306.

Quandt, R. E. and Baumol, W. J. (1966), 'The demand for abstract transport modes: Theory and measurement', *Journal of Regional Science* **6**(2), 13–26.

Radnoti, G. (2002), *Profit Strategies for Air Transportation*, McGraw-Hill. ISBN: 0071385053.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. and Barabási, A.-L. (2002), 'Hierarchical organization of modularity in metabolic networks', *Science* **297**, 1551–1555.

Ravindran, A. R., ed. (2007), *Operations Research and Management Science Handbook*, CRC Press. ISBN-13: 978-0849397219.

Reynolds-Feighan, A. (2001), 'Traffic distribution in low-cost and full-service carrier networks in the US air transportation market', *Journal of Air Transport Management* **7**(5), 265–275.

Reynolds-Feighan, A. (2004), Competition between airline networks, North American Regional Science Association Annual Meeting, Seattle, Washington.

Rosenberger, J. M., Schaefer, A. J., Goldsman, D., Johnson, E. L., Kleywegt, A. J. and Nemhauser, G. L. (2002), 'A stochastic model of airline operations', *Transportation Science* **36**(4), 357–377.

Roth, K. and Miraflor, R. (2004), 'ACES fact sheet', NASA Ames Research Center.

Schleicher, D., Wendel, E. and Huang, A. (2007), Demand loading analysis for a 3X NextGen future, number AIAA 2007-7714.

Shannon, C. E. (1948), 'A mathematical theory of communication', *The Bell System Technical Journal* **27**, 379–423, 623–656.

Simon, H. A. (1955), 'On a class of skew distribution functions', *Biometrika* **42**, 425–440.

Smith, J. C. and Dollyhigh, S. M. (2004), Future air traffic growth and schedule model users guide, Technical Report NASA/CR-2004-213027, NASA Langley Research Center.

Strogatz, S. H. (2001), 'Exploring complex networks', *Nature* **410 (6825)**, 268–276.

Taylor, C. and de Weck, O. L. (2007), 'Coupled vehicle design and network flow optimization for air transportation systems', *Journal of Aircraft* **44**, 1478–1486.

Teodorović, D. (1988), *Airline Operations Research*, Gordon and Breach Science Publishers. ISBN: 2881246729.

The Boeing Company (2008), 'Current market outlook 2008-2027', Boeing Commercial Airplanes Market Analysis, Seattle, Washington.

The Federal Transportation Advisory Group (2001), 'Vision 2050: An integrated national transportation system'.

Trani, A. A., Baik, H., Swingle, H. and Ashiabor, S. (2003), 'An integrated model to study the Small Aircraft Transportation System (SATS)', *Transportation Research Record* **1850**, 1–10.

Transportation Research Board (2001), Aviation gridlock: Understanding the options and seeking solutions, E-CIRCULAR E-C029, National Research Council. Available from `http://onlinepubs.trb.org/onlinepubs/circulars/ec029.pdf`, last accessed May 2009.

Tu, Y. (2000), 'How robust is the internet?', *Nature* **406**(6794), 353–354.

U.S. General Accounting Office (1996), Airline deregulation: Changes in airfares, service, and safety at small, medium-sized, and large communities, Technical Report GAO/RCED-96-79. Available from `http://www.gao.gov/archive/1996/rc96079.pdf`, last accessed May 2009.

Viken, J., Dollyhigh, S., Smith, J., Trani, A., Baik, H., Hinze, N. and Ashiabor, S. (2006*a*), NAS demand predictions, Transportation Systems Analysis Model (TSAM) compared with other forecasts, number AIAA 2006-7761, AIAA Aviation Technology, Integration and Operations Conference.

Viken, J., Dollyhigh, S., Smith, J., Trani, A., Baik, H., Hinze, N. and Ashiabor, S. (2006*b*), Utilizing traveler demand modeling to predict future commercial flight schedules in the NAS, number AIAA 2006-7032, AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference.

Watts, D. J. and Strogatz, S. H. (1998), 'Collective dynamics of 'small-world' networks', *Nature* **393**(6684), 440–442.

Weidner, T. J. (1996), 'Hubbing in U.S. air transportation system: Economic approach', *Transportation Research Record* **1562**, 28–37.

Yan, S. and Tseng, C.-H. (2002), 'A passenger demand model for airline flight scheduling and fleet routing', *Computers and Operations Research* **29**(11), 1559 – 1581.

Yang, L. and Kornfeld, R. (2003), Examination of the hub-and-spoke network: A case example using overnight package delivery, number AIAA 2003-1334, AIAA Aerospace Sciences Meeting and Exhibit.

Zelinski, S. J. (2005), Validating the Airspace Concept Evaluation System using real world data, number AIAA 2005-6491, AIAA Modeling and Simulation Technologies Conference and Exhibit.

Eunsuk Yang was born in Daejeon, Korea to Mr. Kiduk Yang and Ms. Sookja Park. He started his undergraduate study as one of the first students in majoring Aerospace Engineering at Chungnam National University when they opened the department in 1992. He joined the Republic of Korea Air Force in September, 1993 to fulfill his mandatory military requirement as a citizen of Korea. During his term in the force, he performed maintenance and operational support to a fleet of search & rescue helicopters (Boeing HH-47D, Chinook). He resumed his study after he was honorably discharged from the Air Force in February, 1996. After graduating with a Bachelors of Science in Aerospace Engineering, he moved to Atlanta, Georgia to start his graduate study at Georgia Institute of Technology. He received a Masters of Science in Aerospace Engineering in December 2001. He continued his study for a Ph.D., during which he was involved with various advanced systems design projects and optimization problems. His Ph.D. research endeavor ended with the development of a design methodology for evolutionary air transportation networks.