2013-01-01

Development of Demand Forecasting Model For Transatlantic Air Transportation

Katerina Strelcova
University of Texas at El Paso, strelkate@gmail.com

DEVELOPMENT OF DEMAND FORECASTING MODEL FOR TRANSATLANTIC AIR TRANSPORTATION

KATEŘINA STŘELCOVÁ
Department of Civil Engineering

APPROVED:

Ruey Long Cheu, Ph.D., Chair

Salvador Hernandez, Ph.D.

doc. Ing. Ladislav Bína, Csc.

prof. Dr. Ing. Miroslav Svítek

Benjamin C. Flores, Ph.D.
Dean of the Graduate School

Dedication

I dedicate this work to my family and friends who gave me support and motivation during my studies, living in the United States and when I needed it the most.

DEVELOPMENT OF DEMAND FORECASTING MODEL FOR TRANSATLANTIC AIR TRANSPORTATION

by

KATEŘINA STŘELCOVÁ, Ing.

THESIS

Presented to the Faculty of the Graduate School of
The University of Texas at El Paso
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE

Department of Civil Engineering
THE UNIVERSITY OF TEXAS AT EL PASO
May 2013

Acknowledgement

I would like to thank to my advisor Dr. Cheu for his patience, guidance and feedbacks during the whole process of writing my thesis. I would also thank my advisor in the Czech Republic, doc. Ing. Ladislav Bína, Csc. for his suggestion, ideas and arranging meeting with marketing manager of Prague Airport. I also thank to Tomas Horak, ATLANTIS program team member, for his help with all documentation needed and especially for his valuable suggestions during the writing of this thesis.

I would also express special thanks to my family for all their support, encouragement and their love, especially my wonderful mum, sister, and brother. Last but not least, I thank to my friends for their support and that I could rely on.

Abstract

This thesis describes the airline planning process, analyzes the demand of the busiest transatlantic routes and establishes a methodology on how to plan new transatlantic routes from Europe to United States of America with Prague Airport as a hub, or secondary airport for these flights. A few of the most used airports in Europe are chosen and flight data in 2011 from these airports to the United States was analyzed. In addition, demographic, traffic and socioeconomic data of selected airports is collected. A demand forecasting model is then developed in steps using correlation analysis, multiple linear regression analysis and doubly constraint gravity model, and compared with annual number of passengers traveling from Prague Airport to selected destinations in the United States (provided by Prague Airport). The suggested methodology and findings of this thesis can be used for different airlines to forecast demand and plan new transatlantic routes.

Table of Contents

Acknowledgement v

Abstract vi

Table of Contents vii

List of Tables ix

List of Figures x

Chapter 1: Introduction 1

1.1 Background 1

1.2 History and Development of Transatlantic Long-haul Flights 3

1.3 Forecasting of Commercial Air Transportation 6

1.4 Problem Statement - Prague Airport without Long-haul Route 9

1.5 Thesis Objective 13

1.6 Thesis Outline 13

Chapter 2: Literature Review 14

2.1 Levels of Decision in Air Transportation/Airline Management 14

2.2 Demand Forecasting 21

2.3 Methodology of the Development of Demand Forecasting Model 28

Chapter 3: Airport Selection & Data Collection 32

3.1 Route Development – Proposed Approach 32

3.2 External Impacts on Passenger Demand 32

3.3 Performance Indicators Used in Air Transport 33

3.4 Airports Selection 35

3.5 Selection of Variables 39

3.6 Data Set 41

3.7 Methodology of Estimation of Passenger Demand from Supply Data 44

3.8 Limitation of Data 48

3.9 Summary 49

Chapter 4: Development of the Demand Forecasting Model 50

4.1 Demand Forecasting Modeling 50

4.2 Data Input 50

4.3 Correlation Analysis 52
4.4 Regression Analysis 53
4.5 Gravity Model 57
4.6 Summary 68

Chapter 5: Implementation – Prague Airport 69
5.1 Gravity Model Application 69
5.3 Aircraft Suggestion 72
5.4 Suggestion of Frequencies for Flights to New York, Los Angeles, and Chicago 74
5.5 Examples of Airline Schedule 75
5.6. Summary 76

Chapter 6: Conclusion 77
6.1 Summary of the Research 77
6.2 Contribution 77
6.3 Limitations 78
6.4 Future Research 78

References 79

Appendix A 85

Vita…… 86

List of Tables

Table 1.1: Traffic flow from 2003-2031according to RPK 8
Table 3.1: Potential market flights from Prague Airport to U.S. in 2012 36
Table 3.2: Final selection of U.S. airports 38
Table 3.3: Final selection of European Airports 38
Table 3.4: Defined variables 41
Table 3.5: Geo-economic data of the origin airports 43
Table 3.6: Geo-economic data of the destination airports 43
Table 3.7: Sample of index method for estimating number of passengers between city-pairs 47
Table 4.1: Data set for correlation and regression analysis 51
Table 4.2: Correlation analysis of variables 52
Table 4.3: Review of airports with only one direct flight to the selected destinations in 2011 58
Table 4.4: The matrix of estimated number of passengers (PAX_{ij}) as gravity model input 58
Table 4.5: The distance matrix (DIS_{ij}) as gravity model input 58
Table 4.6: The matrix of estimated number of passengers (PAX_{ij}) and summarization of passengers trips 61
Table 4.7: Modification of data by utilization of coefficient 62
Table 4.8: Error results of with different values of x for POP_i and BUS_j 63
Table 4.9: Average absolute percentages errors results of simulation of 10 iterations with different values of x for POP_i and POP_j 65
Table 4.10: Modeling two variables: POP_i and BUS_j with $x=0.8, k=2,3,4$ 66
Table 4.11: Error matric of iteration $k=2,3,4$ and $x=0.8$ 67
Table 5.1: Modeling of configuration: variable BUS_j and POP_i, $x=0.8, k=2$ 69
Table 5.2: Passenger demand of Prague Airport calculated with configuration: a_{min}, $x=0.8, k=2$ 70
Table 5.3: Passenger demand of Prague Airport calculated with configuration: $a_{average}$, $x=0.8, k=2$ 70
Table 5.4: Passenger demand of Prague Airport calculated with configuration: a_{max}, $x=0.8, k=2$ 70
Table 5.6: Results of demand forecasting from Prague Airport 71
Table 5.8: Configuration of A330-200 and A330-300 72
Table 5.9: Current numbers of operated, ordered, and delivered A330-300 in March 2013 73
Table 5.10: Examples of load factors of European airlines and their LF 74
Table 5.11: Suggestion of frequency of flights 75

List of Figures

Figure 1.1: Annual growth of passenger traffic 1990-2011 1
Figure 1.2: World passenger traffic growth & negative impacts on the commercial aviation 1950-2012..2
Figure 1.3: The main points in the history of Czech Airlines connected to transatlantic flights 5
Figure 1.4: World passenger traffic to 2025 (Scheduled services)........6
Figure 1.5: Catchment area of Prague Airport........9
Figure 1.6: Map of current long-haul routes from Prague, Winter 2012/2013........11
Figure 1.7: Current and potential network of long-haul routes in hub-and-spoke network with utilization of Prague such as hub 12
Figure 2.1: Type of decisions 15
Figure 2.2: Steps of airline schedule planning........17
Figure 2.3: Process of airline planning 18
Figure 2.4: Example of a crew trippair 20
Figure 2.5: Typical components of the airline revenue management (RM) system........21
Figure 2.6: Classification of forecasting techniques........23
Figure 2.7: Example of hypothetical correlation between independent variable (gross national product) and dependent variable (number of active aircraft) 25
Figure 2.8: Example of gravity model figure of international air passenger network structure among Asia, America, and Europe in 2000........27
Figure 2.9: Sequence of steps to develop model for long-haul routes........29
Figure 3.1: Relationships among the common passenger indicators........34
Figure 3.2: The busiest airport by passenger in terms of number of passengers in 2011........35
Figure 3.3: Potential market flights from PRG to U.S. in 2012 37
Figure 3.4: Top Transfer points PRG-North America to U.S. in 2011 37
Figure 3.5: Process of selection variables for modeling........40
Figure 3.6: Types of data 42
Figure 3.7: Process of estimation of passenger demand from the supply data........45
Figure 3.8: Number of international passengers in 2012 for calculation of the indexes 48
Figure 4.1: Graph of average absolute percentages errors for POP_i and BUS_j........64
Figure 4.2: Graph of error results of simulation of 10 iterations for POP_i and POP_j 65
Figure 5.1: Comparison of the result of the gravity model and available MIDT data 71
Figure 5.2: Dimensions of A330-300 74

Chapter 1: Introduction

1.1 BACKGROUND

Today, commercial air transportation plays an important role in connecting people and businesses in the world. The aviation industry needs to be improved, not just in safety and security, but in the planning and services too, so that operations can be more efficient and profitable. Passengers are demanding airlines to operate on the most direct routes, with more comfortable aircrafts and inexpensive tickets. The airlines should follow these trends and try to enhance passenger's experience and accomplish high level of passenger satisfaction.

Commercial aviation has grown significantly over the last few decades (Abdelghany 2009). According the statistics of International Civil Aviation Organization or ICAO (2013b), in terms of Revenue Passenger Kilometer (RPK) and Gross Domestic Product (GDP), the world's GDP grew at 2.8 percent per year since 1995, while the world's RPK increased at an average annual growth rate of 5.0 percent, which can be observed in Figure 1.1.

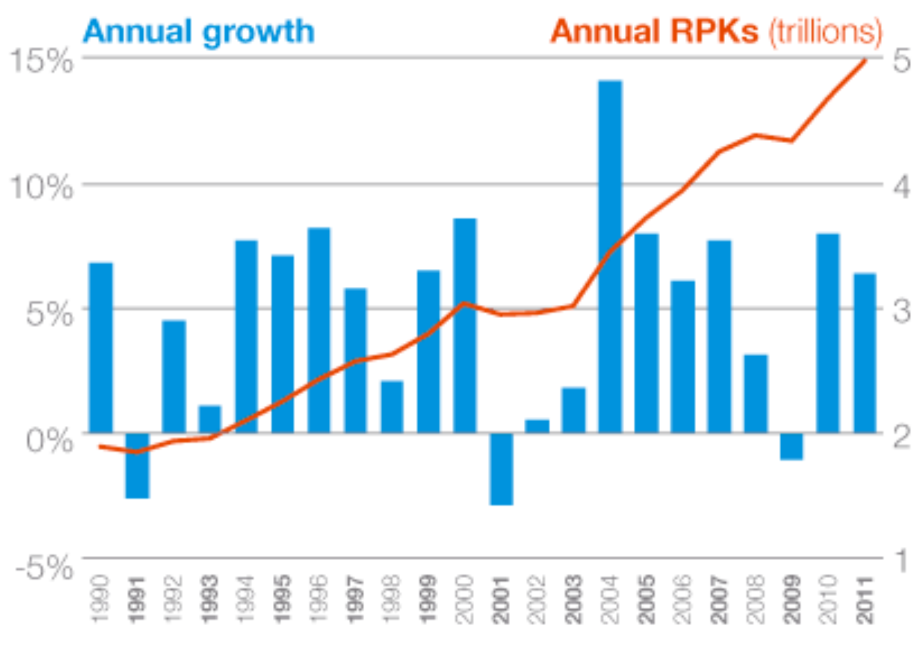


Figure 1.1: Annual growth of passenger traffic 1990-2011

Source: (Boeing 2013)

As can be seen in Figure 1.1, air transportation is a long-term growing industry that expands faster than GDP. The aviation industry is an integral part of the global economy.

Since 1970, as Figure 1.2 demonstrates, the commercial air transport industry has grown at a remarkably steady annual rate of 5% (in terms of RPK), even with the occurrence of multiple wars, oil crisis, financial meltdown, global airline deregulation, deep recessions, the terrorist attacks on September 11, 2001, and other terrorist threats (Boeing 2012). Figure 1.2 also shows that air traffic has doubled every 15 years since 1970.

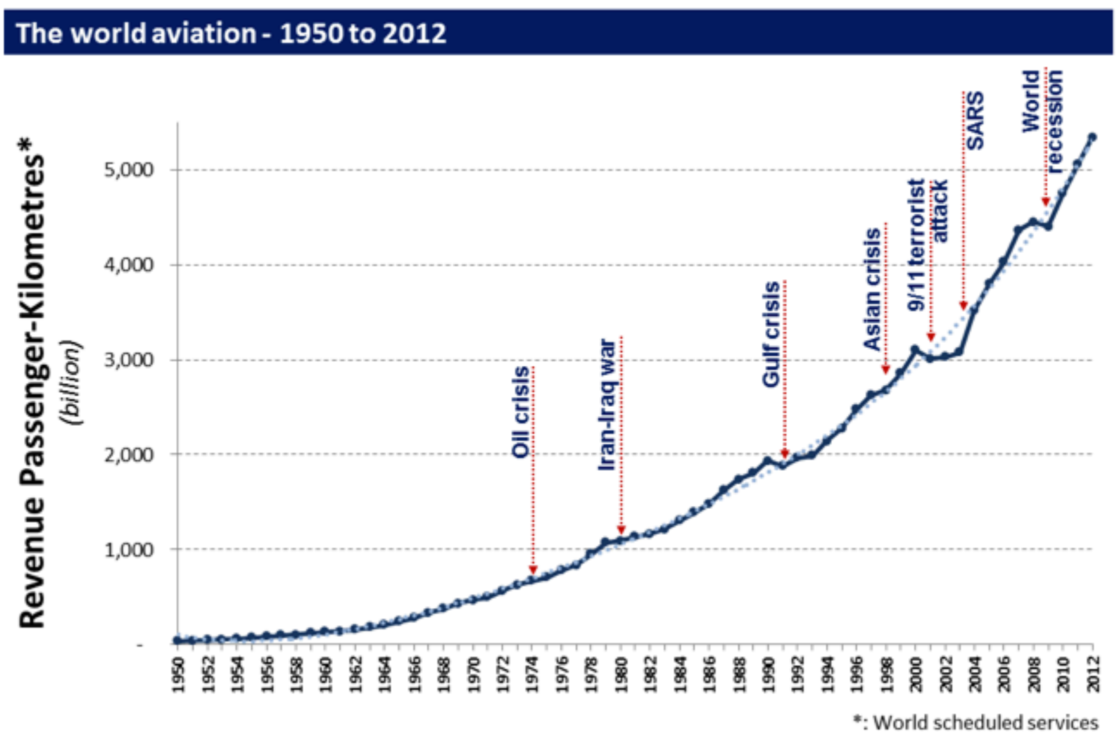


Figure 1.2: World passenger traffic growth & negative impacts on the commercial aviation 1950-2012

Source: (ICAO 2007)

On the other hand, the air transportation industry is really dynamic. There are many events that create positive and negative impacts, which always affect profitability and commercial development of airlines and airports throughout the world.

For example, the merging of airlines, technological development, development of infrastructure and Air Traffic Management (ATM) changes can be viewed as positive external factors. In contrast,

capacity limits and airport congestion increase in fuel prices, and political instability can negatively influence air transportation. More explanation is provided in next chapter.

Currently, there are 240 commercial airlines in 160 countries registered in the International Air Transport Association or IATA (IATA 2012). The increasing number of commercial airlines has constantly forced airline management to perform efficiently, reduce cost, and yet increase revenue in the competitive environment. Furthermore, the increasing demand for commercial air transportation in certain markets has given airlines new opportunities to grow (Abdelghany 2009).

In summary, air transportation is considered to be one big complex system, connecting multiple players, including airlines, airport operators, passengers, government regulators, international associations, and etc. Route forecasting is one of the numerous decisions made by airlines and it represents a critical part of a profitable network planning. Especially, the identification and forecast of new market and revenue can potentially lead to an increase in profit for the airline. This could lead to increased passenger demand from the new market.

Airline route decision can be made by an individual's judgment based on experience, but it becomes more challenging when the number of routes and size of the airline increase.

1.2 History and Development of Transatlantic Long-haul Flights

The history of non-stop transatlantic flights started at 7:52 a.m., May 20, 1927 when an aircraft named "Spirit of St Louis" piloted by Charles Lindbergh took off from Long Island, New York (Michikian 2002). He completed the flight on May 21, 1927, at 10:24 p.m., by landing at Le Bourget Airport in Paris. This first transatlantic flight took 33 and half-hours and covered the distance was 3,500 miles (Daniels 2013; Michikian 2002)

Crossing the Atlantic was connected with a lot of difficulties, such as changing weather, long distance, competition from marine transportation, lack of grand landing rights and so on. Meanwhile, crossing the Atlantic was made by Zeppelin or by some European companies that had made experimental flights carrying only mails. Then finally, the commercial development of the passenger air transportation between North America and Europe started in 1939 when the first commercial transatlantic passenger flights, the Pan American "Yankee Clipper" flying passengers between New York and Marseilles, France, or Southampton, England. The flight took about 29 hours, and cost $375 (that is $5,188, in 2007 dollars) one way (Carryer 2007). All commercial flights were operated with

piston engine powered airplanes until 1952. The first transatlantic commercial jet operation with DH Comet 4 was realized in 1958. Jet engines powered aircrafts slowly replaced piston engine powered aircrafts in 1960. Other development included the introduction of twin engine aircraft (A310, B767), powered with modern fuel-efficient turbofan engines in 1980s (European Academy For Aviation Safety 1999)

1.2.1 History of Transatlantic flights from Prague

As Francis et al. (2007) mention in his study, the definition of long haul route was all transatlantic flights according to Association of European Airlines (AEA). "Hence, the cut between medium-haul and long-haul is around six hours of flying time. The Sahara Desert and the Atlantic Ocean form natural clear cut dividers when travelling from Europe" (Francis et al. 2007).

One of the official manuals of ICAO (2011) says, there is no clear definition of long-haul routes and defined long-haul routes in relation to the aircraft type. "The long-haul group would include aircraft capable of a maximum range of more than 8,000 km (e.g. A330, A340, A380, B747, B767-200ER, B763, B764, B777, IL96)" (ICAO 2011).

For the purposes of this thesis the term transatlantic flights are defined as flights connecting countries on both sides of the Atlantic, typically between EU countries and U.S.

Transatlantic long-haul routes have been an important part of air transportation since the beginning of commercial aviation in 1939, when the flights were regularly flown from Europe to North America, South America, Africa, the Far East and Australia, and vice versa. Czech Airlines, the national carrier of the Czech Republic, operated the first transatlantic flight from the Czech Republic from Prague to Havana in 1962 (Czech Airlines 2013).

Czech Airlines was founded as Czechoslovak State Airlines (CSA) on 6th of October 1923. The first commercial flight was from Prague to Bratislava and was held in the first month of foundation. In 1995, Czechoslovak State Airlines was renamed as Czech Airlines (AirlineUpdate.com 2013). The CSA acronym is still used by Czech Airlines today.

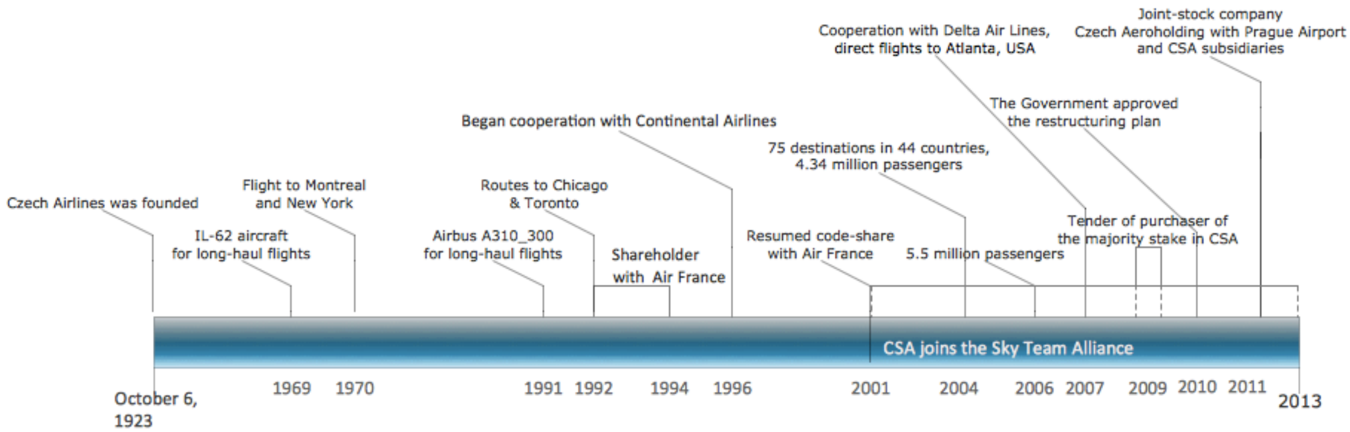


Figure 1.3: The main points in the history of Czech Airlines connected to transatlantic flights

adapted from (Czech Airlines 2013)

1.2.2 History of the Czech Airlines fleet

“The airline's first transatlantic flight, with the leased Bristol Britannia BB-318 aircraft, took place in 1962, flying from Prague to Havana” (Airliners.net 2008). From Figure 1.2, IL-62 started to be used in 1969 as the first aircraft for long-haul routes. Then, Airbus 310 replaced IL-62 in 1991. Czech airlines operated four aircrafts in total; other aircrafts were purchased in 2003 and 2004. The CSA fleet consisted of 49 aircrafts in the beginning of 2010. From then on, the number of aircrafts together with operating flights was continuously decreasing with 38 aircrafts in 2011 and 26 aircrafts in 2012. Likewise, the last Airbus A310, the largest aircraft in the CSA fleet, left after 19 years of service to be painted into United Airways Ltd colors within the program of airline restructuring (Czech Airlines 2013, Krupka 2010). Since September 2010, the new joint-stock company, Czech Aeroholding, including CSA, her subsidiaries and Vaclav Havel Prague Airport[1] was created. After one year, the restructuring project was finished finally by registration in the Commercial Registry of the Czech Republic in 2011.

CSA reduced its fleet to 38 jets in the beginning of 2010 and cut flights with a higher share of transfer passengers (Ponikelska 2011). The reduction of flights, sale of planes, surrender of landing slots at European airports are CSA’s responses to continuing consequences of economic recession (CAPA 2012; Eurofound 2009). Nowadays, the fleet includes 26 aircrafts and just 15 (nine A319-100, four A320-200, two A321-200) suitable for medium- and long-haul flights (Czech Airlines 2013).

[1] In this thesis Vaclav Havel Airport Prague is called Prague Airport or PRG

1.3 FORECASTING OF COMMERCIAL AIR TRANSPORTATION

According to ICAO (2012), almost 3 billion people used air transport in 2012, which is up by 5 percent since 2011. This figure is expected to reach at least over 6 billion by 2030. The lowest annual growth rate is estimated at 2.6 percent, the most likely growth rate is 4.6 percent, while the highest growth rate is 6.2 percent (ICAO 2013). Correspondingly, Boeing (2013) expects world airline traffic to grow by 5 percent annually from 2012 to 2031 in terms of RPK. During the same period, Boeing expects the number of passengers to grow by 4 percent per year, and global GDP to grow by 3.2 percent per year.

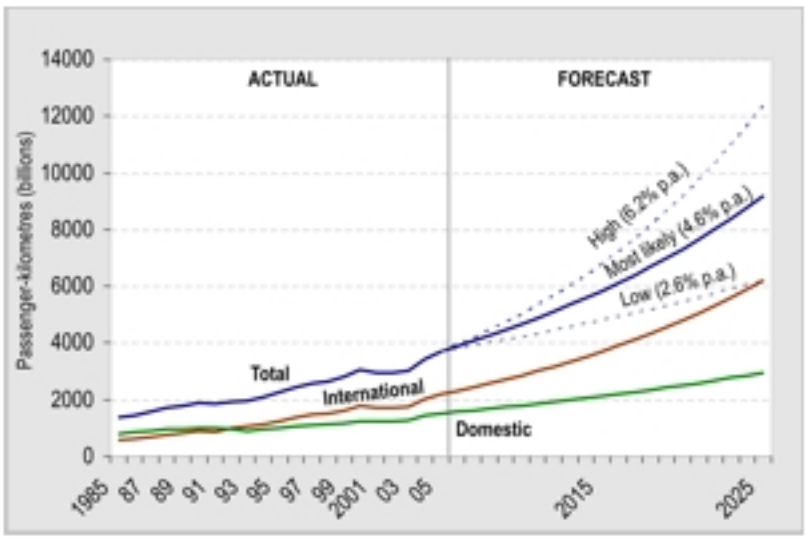


Figure 1.4: World passenger traffic to 2025 (Scheduled services)

Source: (ICAO 2013a)

These above forecasts are based on positive economic development worldwide, despite slow economic growth in some regions and some downturns in the past, and the implementation of severe fiscal policies in key European economies (ICAO News 2012).

As shown in Figure 1.4, international traffic grew by 6.5 percent in 2012, the same rate as the previous year. North America registered the lowest growth rate of all international markets, but it is still has positive growth (ICAO News 2012).

Boeing (2012) forecasted that the growth of air transportation is expected to continue over the next 20 years. These predictions are also supported by FAA (2012), which expected that passenger demand will almost double in the next 20 years.

Table 1.1: Traffic flow from 2003-2031according to RPK

adapted from (Boeing 2013).

Traffic flow	RPK in billions										Annual growth (%)
	2003	2004	2005	2006	2007	2008	2009	2010	2011	2031	2011 to 2031
Europe to and from Middle East	72	80	87	99	107	115	131	144	153	417	5.1
Europe to and from North America	**349**	**376**	**391**	**403**	**421**	**432**	**405**	**419**	**430**	**901**	**3.8**
Europe to and from Northeast Asia	54	58	58	59	65	66	57	62	62	124	3.5
Europe to and from South America	48	55	64	67	71	75	79	83	90	216	4.5
Europe to and from South Asia	33	38	43	53	59	55	51	54	54	227	7.4
Europe to and from Southeast Asia	92	101	100	98	92	100	98	100	104	278	5.0
World Total	**3,304**	**3,754**	**4,026**	**4,234**	**4,539**	**4,611**	**4,519**	**4,885**	**5,198**	**13,764**	**5.0**

Table 1.1 shows the slow growth of transatlantic market from Europe to North America. It is probably because of the combined effect of European debt crises and the economic recession in U.S. In contrast, air travel from Europe to East Asia and Middle East has the highest growths.

1.4 Problem Statement - Prague Airport without Long-haul Route

This thesis focuses on the transatlantic flights from Prague to the United States (U.S.). Prague Airport is a modern airport with a strategic position in the middle of Europe. It has the potential to be one of Europe's international air hubs for its location and the contemporary equipment. It has been a part of the joint stock company, Czech Aeroholding, since 2011. Prague Airport is already an important hub for Central and Eastern Europe, and serves almost 12 million passengers annually. It is also the biggest airport among the new EU member states, with the current capacity of 15.5 million passengers per year. The airport has a catchment area with 2.5 million people living within 60 minutes and 8 million people living within 120 minutes, as shown in Figure 1.5. This adds to the possibility of attracting more passengers (Prague Airport 2013b: Prague Airport 2013c).

Figure 1.5: Catchment area of Prague Airport

source: (Prague Airport 2013c)

Prague airport has won various prestigious awards, for examples the 2011 Eagle Award by IATA for being the best developing airport and the Award for Best Airport in Eastern Europe in the 2008 World Airport Awards (Prague Airport 2013b).

Fifty scheduled airlines operated direct flights from Prague Airport to 120 destinations in 46 countries in 2011 (Prague Airport 2012). However, these numbers are decreasing (Prague Airport 2013b). In the past, CSA and Delta Airlines operated direct flights to New York. In 2009, CSA cancelled this route because it was not profitable, even though this route had a high load factor. Delta Airlines continued to serve this route until November 18, 2012. Delta Airline's reason for dropping the route was due to high fuel cost and fluctuating passenger demand. According to the information of marketing director of Prague Airport, the load factor of that route was almost 99 percent in spring 2012 (Skalicky 2012).

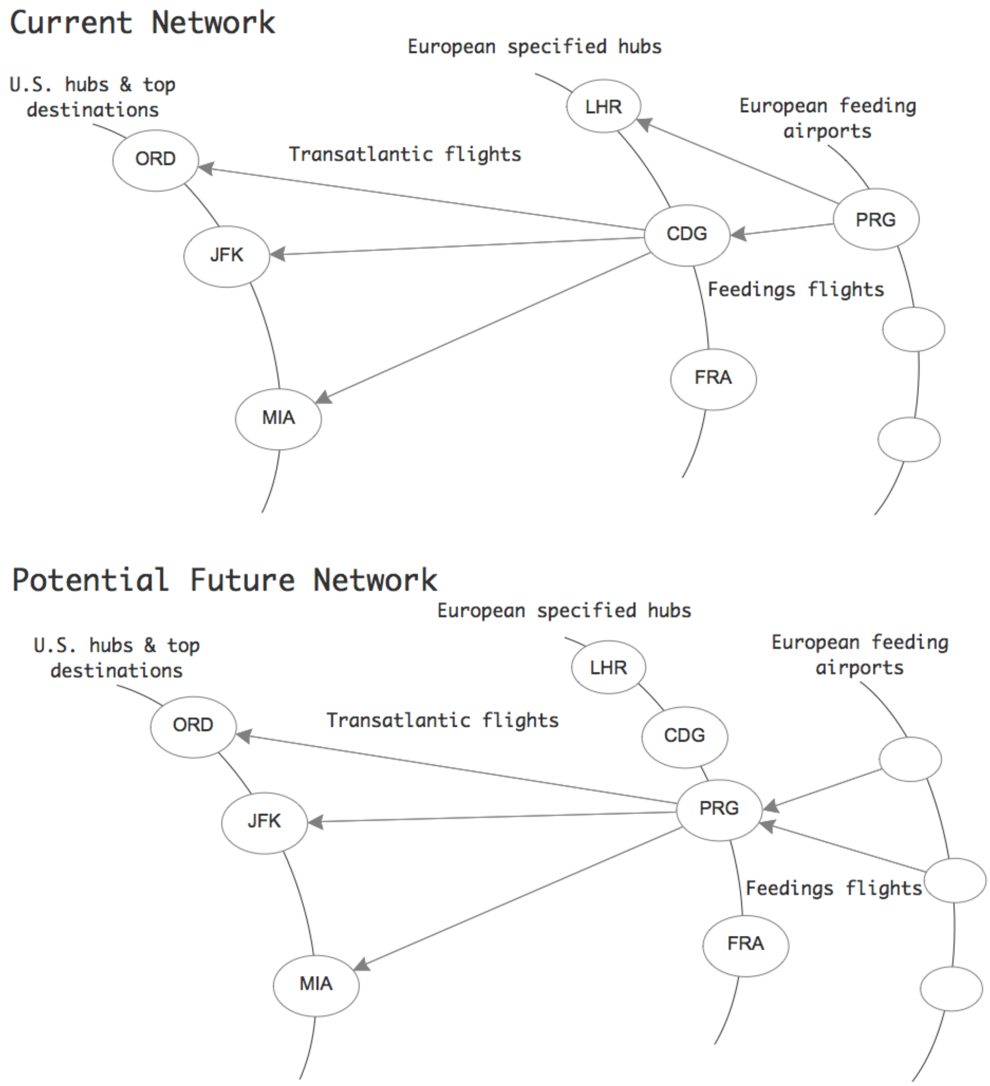
Additionally, the economic problems of CSA have forced it to sell all the aircrafts for long-haul flights. By canceling the last direct flight to U.S., CSA ceased to serve transatlantic routes. Currently, there is no way to get to U.S. directly from Prague by any commercial scheduled flight. Passengers traveling from Prague to U.S. airports must make a flight connection at other European airports.

In summary, Prague Airport has current handling capacity of 15.5 million passengers per year, modern equipment and the very strategic position. According to PRG (2013a), "During the Winter 2012/2013 timetable (valid from October 28, 2012 to March 30, 2013), 46 scheduled airlines offered direct flights from Prague Airport to 83 destinations in 38 countries." There is no direct flight to U.S. The new plan of Czech Airlines is to offer new direct route to Seoul by using A330. The planned long-haul flights are targeting at airports in East Asia.

Figure 1.6: Map of current long-haul routes from Prague, Winter 2012/2013

Source: (Prague Airport 2013a)

The favorable conditions of Prague Airport might inspire the resumption of transatlantic flights. Therefore, there is a need to conduct research to try to determine if there is enough transatlantic passenger demand to and from Prague. The thesis might help to Prague Airport to establish a future transatlantic network structure or at least, bring new idea how to model long-haul routes from Prague to U.S.



ORD … Chicago O'Hare International Airport, JFK … John F. Kennedy International Airport, MIA … Miami International Airport, LHR … London Heathrow, CDG … Charles De Gaulle Airport, PRG … Vaclav Havel Airport Prague, FRA … Frankfurt Airport

Figure 1.7: Current and potential network of long-haul routes in hub-and-spoke network with utilization of Prague such as hub

Figure 1.7 illustrates current situation of long-haul routes from Europe to U.S. and the placement of Vaclav Havel Airport Prague in this network. Currently, Prague Airport is used mostly as feeding airport of larger European airports such as London Heathrow, Paris Charles De Gaulle, or Frankfurt Airport, for long-haul routes to U.S. This thesis suggests using Prague Airport as a hub or secondary airport (that receives feeders) for long-haul route instead of just feeding the long-haul routes to its competitors. The bottom part of the sketch shows the scenario where Prague Airport occupies the

strategic position in the transatlantic air transport network and it becomes a potential competitor of current European hubs.

1.5 Thesis Objective

The aim of this thesis is to develop a transatlantic air passengers forecasting model for airlines to plan direct flights from Prague to selected U.S. cities.

The thesis will describe the possibilities of collecting historical and existing data, analysis methodology, which can help to identify the main factors that influence the passenger demand of the transatlantic route network.

This model should help airlines operating at Prague Airport to plan direct transatlantic long-haul routes to U.S. cities. Although this model describes the solution for Prague Airport and U.S. route network. It should be general enough for planning routes in similar environment with similar conditions (e.g., from airports in central and eastern Europe to U.S.) and bring new ideas for the modeling of long-haul routes.

1.6 Thesis Outline

Chapter 1 focuses on the background of this thesis, history of transatlantic flights, history of Czech Airlines, problem statements and motivation for this research.

Chapter 2 primarily reviews the airline schedule planning process and methods for air passenger demand forecasting.

Chapter 3 explains the methodology used in this thesis in the development of the model for new transatlantic routes, including demand forecasting model, aircraft selection, frequency of service and flight schedule. This chapter also includes a description of the data collected.

Chapter 4 describes the development of the demand forecasting model for transatlantic air passenger transportation.

Chapter 5 shows the implementation of the demand forecast modeling for Prague Airport.

Chapter 6 summarizes important findings of the thesis and gives recommendations based on the current study.

Chapter 2: Literature Review

This chapter is structured in two main sections. The first section introduces decision levels in airline management followed by an overview of the network structure and airline schedule planning process. The second section focuses on demand forecasting including the techniques and models to use for estimation.

2.1 LEVELS OF DECISION IN AIR TRANSPORTATION/AIRLINE MANAGEMENT

Airline management consists of three layers of decision, strategic, planning and operation, which are summarized in Figure 2.1. All these decisions are interrelated. For example, strong demand forecasting might call for a change in a strategic and also in planning decision regarding expansion or increase and decrease in fleet size, change in number of flights, or change in origin/destination (Abdelghany 2009).

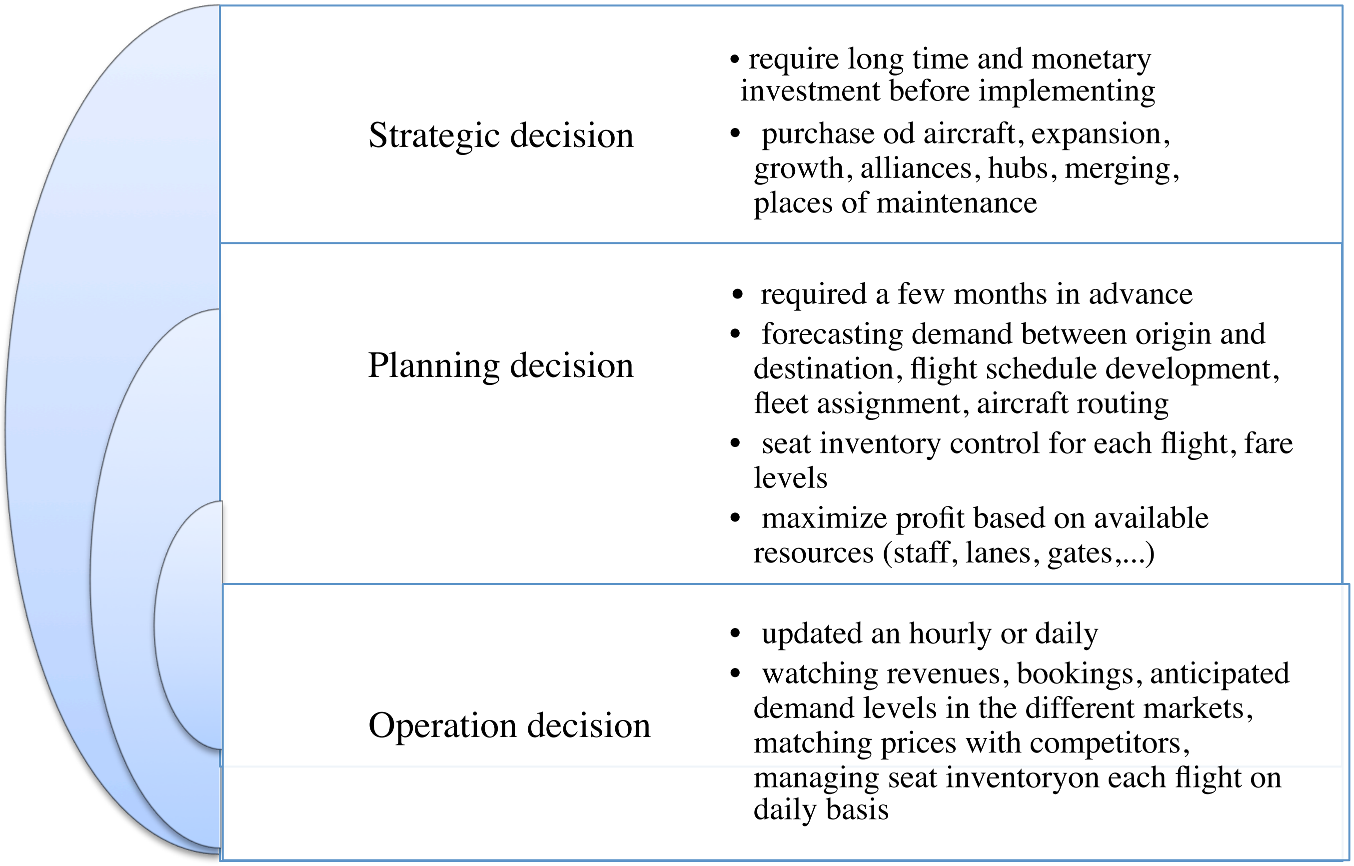


Figure 2.1: Type of decisions

adapted from (Abdelghany 2009)

2.1.1 Network Structure as a Strategic Decision

Most of the systems of airline´s flights are based on one of the two systems, linear network and hub network, or their combination.

Most of the commercial airlines are scheduled airlines with predefined flight schedule including (i) the specification of the airports they flies to; (ii) departure times; (iii) capacity (number of seats) of each flight in the schedule and (iv) predefined network structure. Very few airlines are charter airlines that are typically operated on demand basis (Abdelghany 2009). This thesis focuses on airline planning

and route development of scheduled airlines operating out of European airports (especially Prague Airport) and serve U.S. market.

The most common network structures are hub-and-spoke network that usually consists of one or more airports in the network as a hub (Abdelghany 2009). Big airlines such as United Airlines, Air France, British Airways and Delta use the hub-and-spoke network structure. Other types of networks can be point-to-point that is defined as network with non-stop flights between all airports (Grosche 2009), or combination of both of the networks. Network structure is one of the major strategic decisions of the airline (Abdelghany 2009).

For example, the research of Hansen & Wei (2006) shows that airlines operating with hub-and-spoke networks are the most influenced by density of the network. This research also proved that if airlines would focus on increasing frequency of flights, rather than increasing size of aircrafts in hub-to-spoke networks, they could interest more customers.

2.1.2 Airline Schedule Planning and Operations

According to Abdelghany (2009), planning starts by recording and anticipated demand and supply and taking into consideration the available airline sources. Next, route development and schedule development is followed, including schedule planning, fleet assignment, aircraft routing, crew scheduling, airport staff scheduling, pricing and seat inventory control, sales and marketing initiatives.

Airline scheduling is a difficult problem that cannot be represented simultaneously and solved as a single problem but has to be divided into sub-problems and be solved in steps, because many decisions in airline schedule planning process have traditionally been classified and optimized in a sequential manner (Grosche 2009; Lohatepanont & Barnhart 2004). The scheduling methodology also depends on airline size and network structure. Big airlines companies usually use special software systems for schedule planning.

As shown in Figure 2.2, the airline schedule planning is decomposed into two main procedures (Lohatepanont & Barnhart 2004):

- Route development,
- Schedule development.

2.1.2.1 Route Development

The route selection decision might be part of strategic and/or planning decisions. Also, the route development is the first step of schedule planning that takes place at least one year before the operation of the flight (Lohatepanont & Barnhart 2004; Abdelghany 2009).

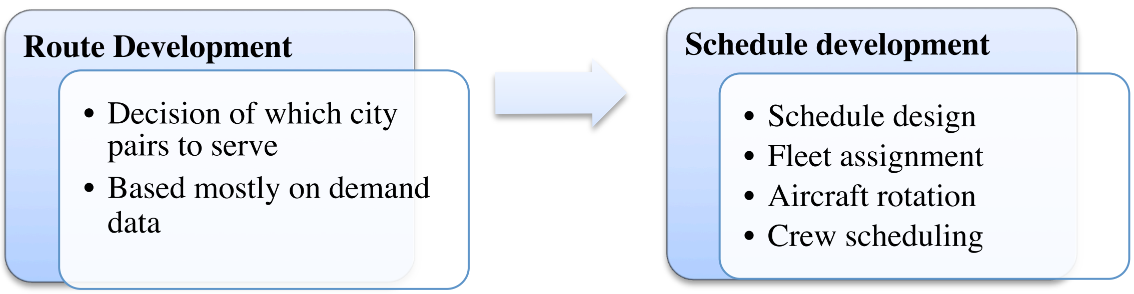


Figure 2.2: Steps of airline schedule planning

adapted from (Lohatepanont & Barnhart 2004)

2.1.2.2 Schedule Development

The airline schedule is planned one year before the operation of a flight. Schedule development mainly includes schedule design, fleet assignment, aircraft routing, and crew scheduling. It is commonly modified from the existing schedule by presenting the changes in demand or environment based on forecasting. From Figure 2.3 is obvious that airline planning is a continuous process and each step involves the inputs from the predecessor and outputs to the successor.

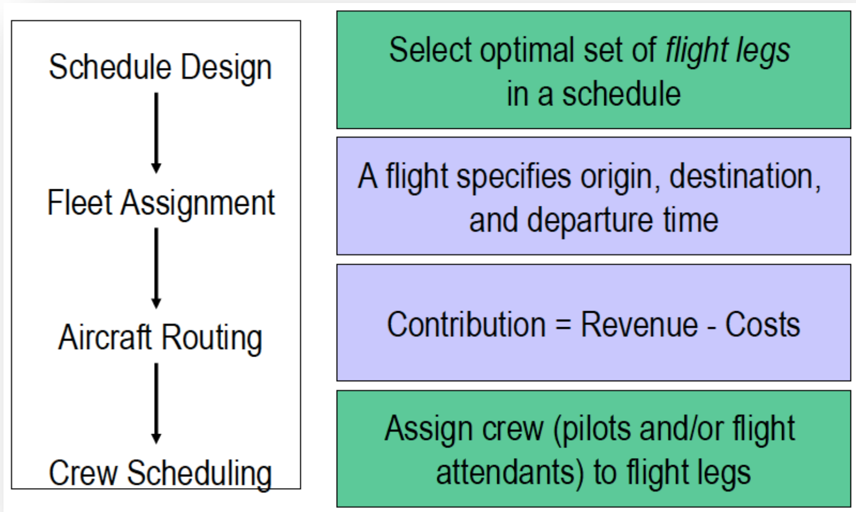


Figure 2.3: Process of airline planning

Source: (Belobaba 2006)

I. Schedule Design

Schedule planning usually comes out from an existing schedule, with a well-developed route network (Lohatepanont & Barnhart 2004).

For example, the study by Lohatepanont & Barnhart (2004) describes how to design flight schedules that bring maximum profit based on the actual flight schedule data from a major U.S. airline from the current and previous seasons. The next season's schedule is developed from previous ones by modification of mandatory and optional flights.

According to Belobaba (2006) and Lohatepanont & Barnhart (2004), schedule design is divided into two sequential steps. First, frequency planning determines how often should service be provided on defined route network between specific origin and destination pairs. More frequent departures can reduce "wait time" between flights, improves convenience for passengers, and increases market share. One of the most important aspects is to capturing time sensitive business travelers. Second, timetable development is usually created from previous and current timetables in exact times 2-6 months in

advance. It has to be in balance with aircraft rotations, network consideration and other constraints. For example, passengers do not want to arrive at their destination before 6am on flights from U.S. to Europe.

II. Fleet Assignment

Fleet assignment is a process that is necessary for airlines operating more than one type of aircraft. The goal of this process is to choose the right type of aircraft to operate the different flights that matches with passenger demands.

Aircraft travel range has to be consistent with the distance of the flight and take into consideration the most optimal fuel consumption. In addition, airport characteristics at the origin and destination of the flight have to be reflected such as adequate aerodrome reference code of the runway, allowed noise level, number of gates, and equipment at the airport (Abdelghany 2009).

Fleet assignment is connected closely to schedule design, especially in case of planning large or small aircrafts interconnected with number of flights in the market.

III. Aircraft Routing

The purpose of aircraft routing is to maintain a feasible rotation of aircrafts with respect to maintenance rules and needs at the right maintenance stations (Lohatepanont & Barnhart 2004).

IV. Crew Assignment

This refers to the monthly schedule for airline crew with predefined trip pairs that are defined as a sequence of flights, which originates and ends at the home city (crew domicile). An example of assigned crew schedule is depicted in Figure 2.4.

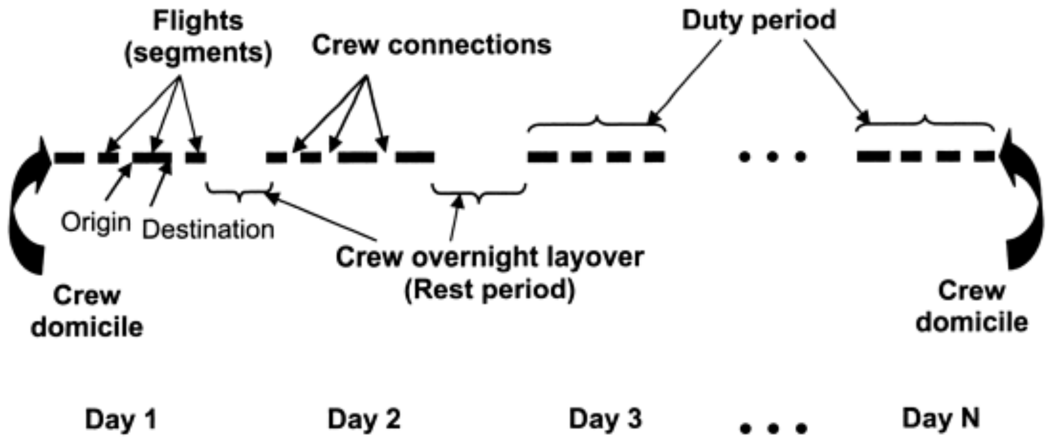


Figure 2.4: Example of a crew trippair

Source: (Abdelghany 2009)

2.1.3 Revenue Management

"Revenue management is defined as selling the right seat to the right customer at the right price and at the right time" (Abdelghany 2009). Likewise, "Revenue management is the subsequent process of determining how many seats to make available at each fare level" (Belobaba et al. 2009).

The revenue management focuses on business passengers and makes sure that there are always enough seats for them, because they bring profit to the airlines. Business passengers usually require flexible tickets and are able to pay more than theleisure passengers. Leisure passengers are more sensitive to the price of the tickets and they have lower budgets for tickets.

Revenue management involves the three main modules (Abdelghany 2009):

- pricing (set up the optimal price including competitors pricing in each specific market)
- demand forecasting
- seat inventory control (assigned seats of each flight to the concrete demand to maximize total revenue)

Revenue management is calculated by mathematical formula that maximizes total revenue considering fare-class, demand, seat capacity, and so on (Abdelghany 2009). There are some pricing and inventory decision support tools that helps airlines to more accurately forecast future demand and set up ticket prices. One of them is Sabre® AirVision™ Revenue Manager (Sabre 2013).

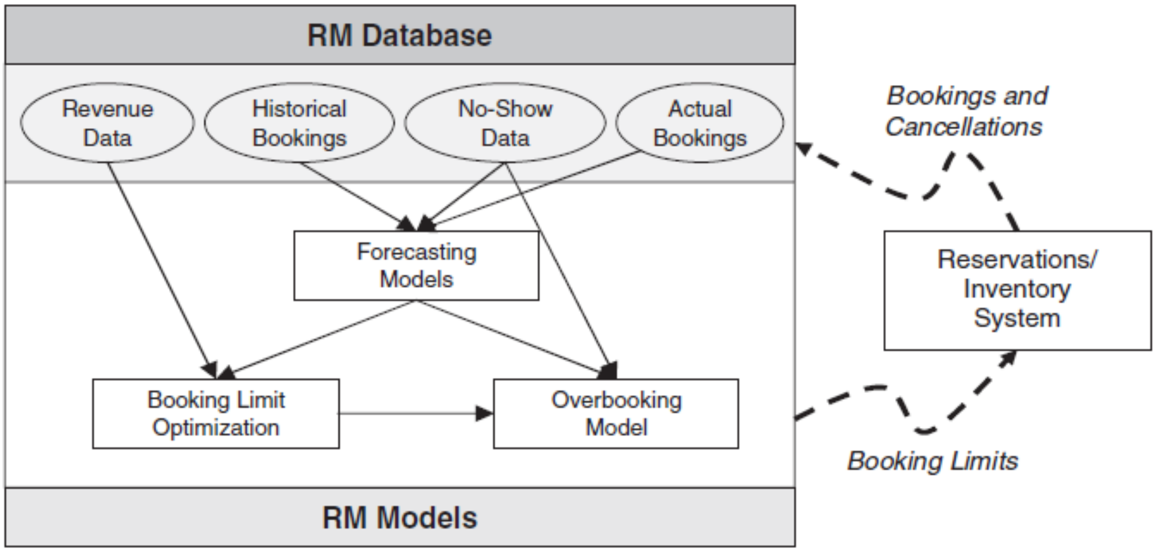


Figure 2.5: Typical components of the airline revenue management (RM) system

Source: (Belobaba et al. 2009)

Figure 2.5 is example of typical components of the airline revenue management (RM) system that is highly based on historical and current data that support forecast models to optimize the booking processes.

2.2 DEMAND FORECASTING

Demand forecasting may be described as the process of predicting and estimating the expected number of passengers (passenger volume) between two airports for a given time interval in the future. In some cases, demand forecasting also includes estimating the possible changes in demand due to the changes in the schedule, pricing, competition, and airline share in the origin-destination market (Abdelghany 2009).

In general, demand forecasting is the main tool for airlines to predict the future behavior of potential passengers so that it can be ready to offer service in the market where it will be needed and profitable. It is usually quantified in terms of currency (revenue) or in revenue passenger miles (RMPs) in a defined period of time (Wensveen 2007).

Based on the demand forecast, airlines can make decisions about opening new routes or optimization of the existing routes. There exist several demand forecasting techniques, but in practice

these different techniques are combined or compared with each other because no single one can guarantee high accuracy (Grosche et al. 2007).

2.2.1 Forecasting Techniques

Forecasting methods exists from very simple to very sophisticated. Figure 2.6 illustrates a classification of forecasting techniques. Forecasting techniques can be divided into two main groups: qualitative and quantitative. The choice of forecasting techniques should be based on several factors, i.e.: what the forecast is going to be used to for, availability of data and its quality and accuracy, possible data processing techniques (Wensveen 2007).

The biggest airport and airlines do not always use the most sophisticated methods. Moreover, complex methods do not always result in better forecasts (TRB 2007). For forecasting, one may use causal models that are based on a statistical relationship between the dependent (forecasted) variable and independent (explanatory) variables, or judgmental methods that are based on assessment of experts (Wensveen 2007). The techniques are described in more details below.

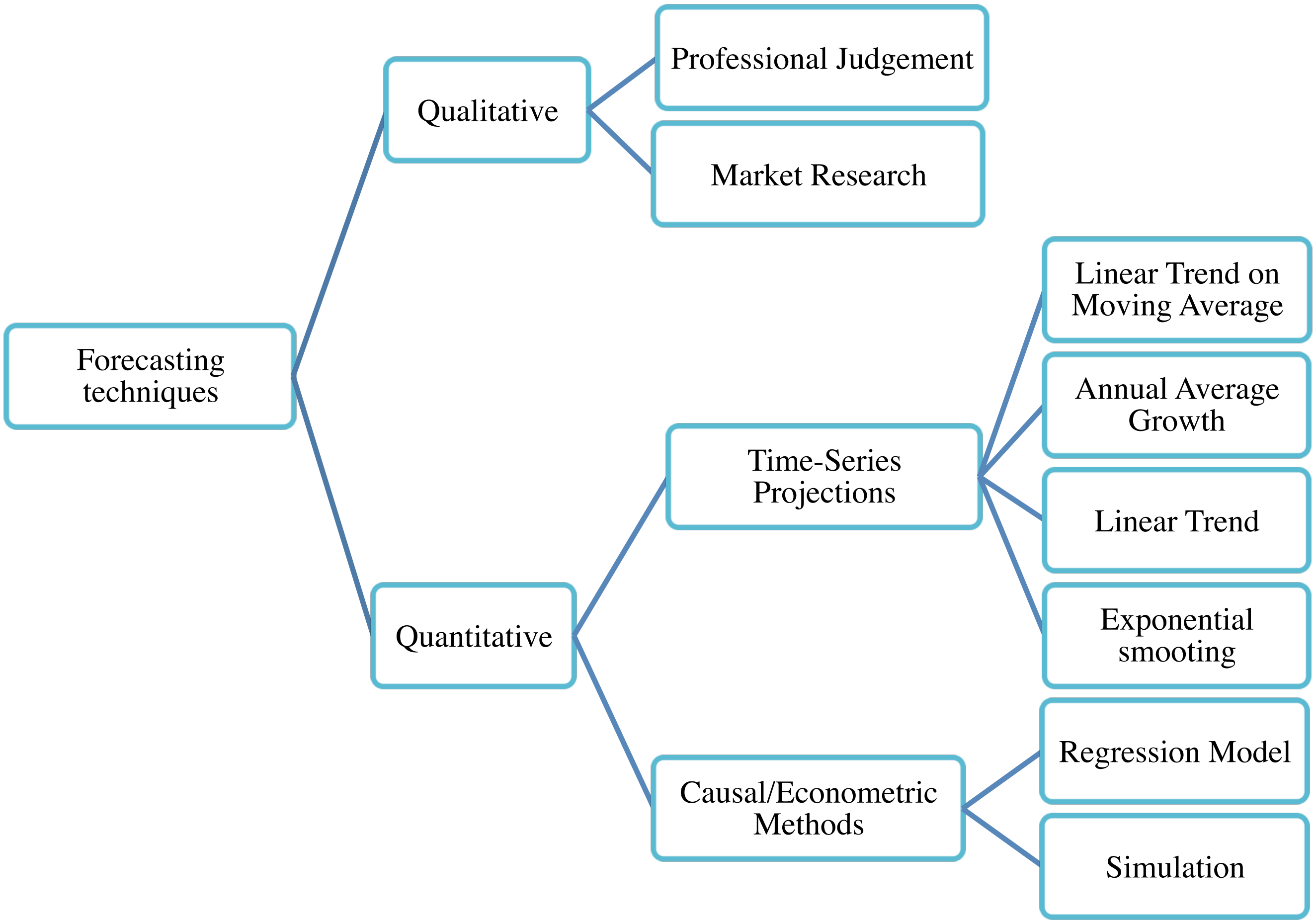


Figure 2.6: Classification of forecasting techniques

Source: (Grosche 2009)

2.2.1.1 Qualitative Techniques

Qualitative techniques are rough guesses, analogy conclusions, and prediction of future trends. Even though, they are not lined up among the most accurate ones, in most of cases, they can be very powerful tools for prediction. Moreover, it may be the only available tool when neither historical data nor other supporting information, or analysis is available.

2.2.1.2 Quantitative Techniques

Quantitative methods are relatively more sophisticated methods that use mathematical models and investigate relations between independent and dependent variables. It allows independent prediction of future trend based on available data and facts.

The most important quantitative methods are (Grosche 2009; Wensveen 2007):

- **Time-series Projections**. These are mathematical models based on statistical correlation that is built from demand data as a dependent variable and time as an independent variable. The historical sequence of data is plotted on a graph and a trend line is established. This method cannot be used for prediction of a new route where the historical data is not available.
- **Causal/Econometric Methods**. These methods express the statistical relationship between air travel demand and selected economic or supply variables that stimulate travel. Causal models are better in describing passenger's behavior. Typically, the most used mathematical model is regression model. One of its alternatives is gravity model (Grosche 2009; Wensveen 2007). For model development the variables have to be selected. It usually depends on the resources available and judgment of the researcher.

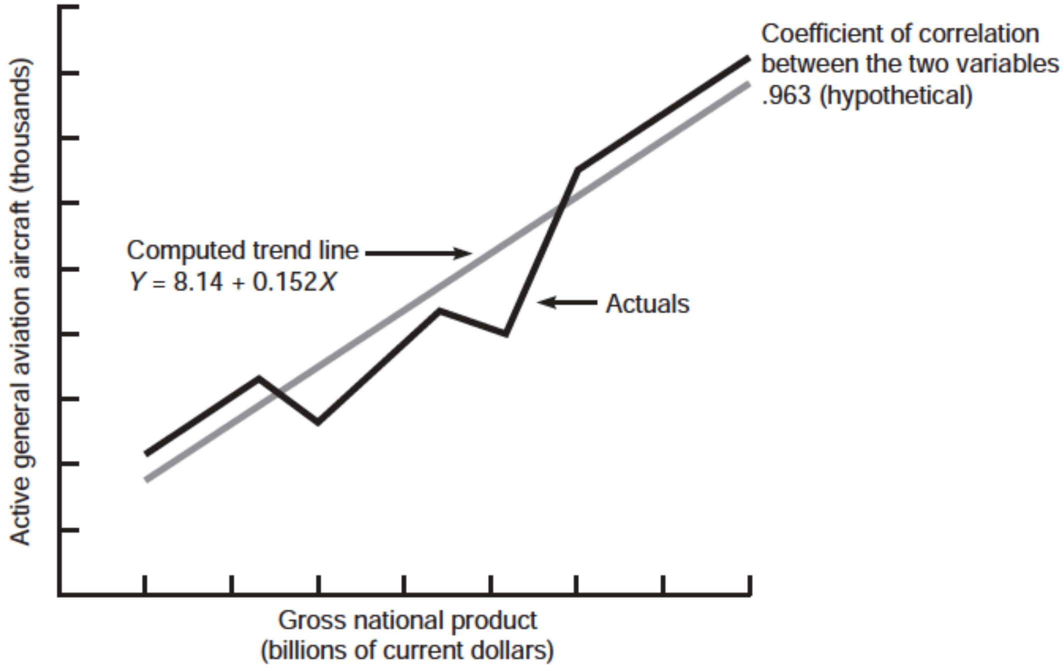


Figure 2.7: Example of hypothetical correlation between independent variable (gross national product) and dependent variable (number of active aircraft)

Source: (Wensveen 2007)

As is shown in Figure 2.7, independent variables used in forecasting the air transportation and usually contains Gross National Product (GNP), or not mentioned in the Figure 2.8, personal income, and spending on services. On the other hand, dependent variables can include values such as revenue passengers enplaned, RPK, passenger revenues, or number of active aircrafts (Wensveen 2007).

2.2.2 Gravity Model

As mentioned before, the most common casual model in air transportation forecasting is the gravity model. Based on the study of (Chang 2011), the general form of the gravity model for analyzing passenger demand can be expressed, for example, as

$$T_{ij} = \frac{CP_i A_j}{f(d_{ij})}, \qquad (2.1)$$

where T_{ij} represents the number of trips produced in country i (origin) and attracted to country j (destination), C is a constant, P_i is the production factors of country i, A_j is the attracting factors of country *j,* and d_{ij} is the distance between country *i* and country *j*.

For example, the study of Chang (2011) showed that GDP and import values have a positive impact on passenger flow including national per capita income for the origin country in regression analysis. In contrast, the unemployment rate has negative influence.

Another formula of gravity model that can be used in trip distribution is (Baik et al. 2006):

$$T_{ij} = k\frac{P_i A_j}{W_{ij}^{c}}, \qquad (2.2)$$

where:

T_{ij} are trips between zones *i* and *j*,

k is a constant,

P_i is a trip production from origin zone *i*,

A_j is attractiveness of the destination zone,

W_{ij} is the impedance between zones *j* and *j*.

The attractiveness variable can be population, number of shopping centers, etc.

(Matsumoto 2007) determined gravity model with using GDP, population and distance as independent variables:

$$T_{ij} = A\frac{(G_i G_j)^{\alpha}(P_i P_j)^{\beta} e^{\delta D_1} e^{\varepsilon D_2} e^{\zeta D_3} \ldots e^{\phi D_{17}} e^{\chi D_{18}} e^{\psi D_{19}}}{(R_{ij})^{\gamma}}, \qquad (2.3)$$

where:

T_{ij} is the net volume of international air passengers over ten thousand between city *i* and *j*

G_i is the real GDP per capita of the country in which city *i* is located, expressed in US dollars at the 2000 exchange rate and constant 1990 prices,

G_j is the real GDP per capita in 2000 of the country, in which city *j* is located, expressed in US dollars at the 2000 exchange rate and constant 1990 prices,

P_i is the population (in thousands) of city *i* in 2000,

P_j is the population (in thousands) of city j in 2000,

R_{ij} is the distance between city i and city j in kilometers,

D is the city-dummy variables,

A is the constant.

After transforming equation into log form, an ordinary least-square regression analysis was used to calibrate the model parameters.

Figure 2.8 represents the example of gravity model figure of international air passenger network by highlighting the biggest volume of air passengers. For example, 8,200 passengers is demand of San Francisco and 10,460 passengers is demand of Chicago from mentioned cities in Figure 2.8.

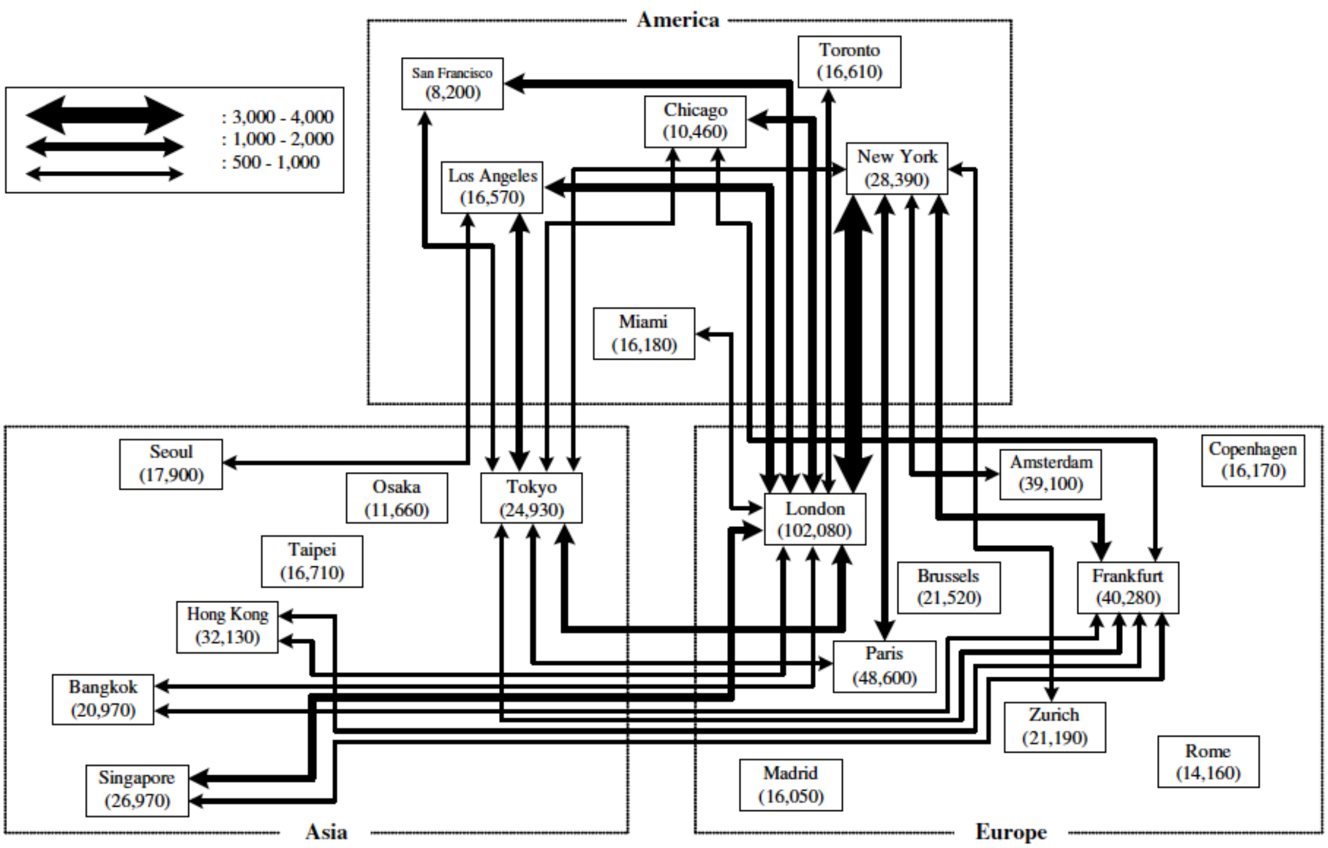


Figure 2.8: Example of gravity model figure of international air passenger network structure among Asia, America, and Europe in 2000

Source: (Matsumoto 2007)

2.2.2.1 Calibration of the Gravity Model

Kanafani (1983) analyzed two types of calibration of gravity model:

- Cross-section – application of the same model for different airport pairs during one time period
- Time-series data – calibration for a specific airport pair using extending data over the period

2.3 METHODOLOGY OF THE DEVELOPMENT OF DEMAND FORECASTING MODEL

This section describes the methodology of route development and airline schedule planning based on the literature review and the author's own ideas.

The development of methodology consists of ten following main steps that are described in more detail in Figure 2.9.

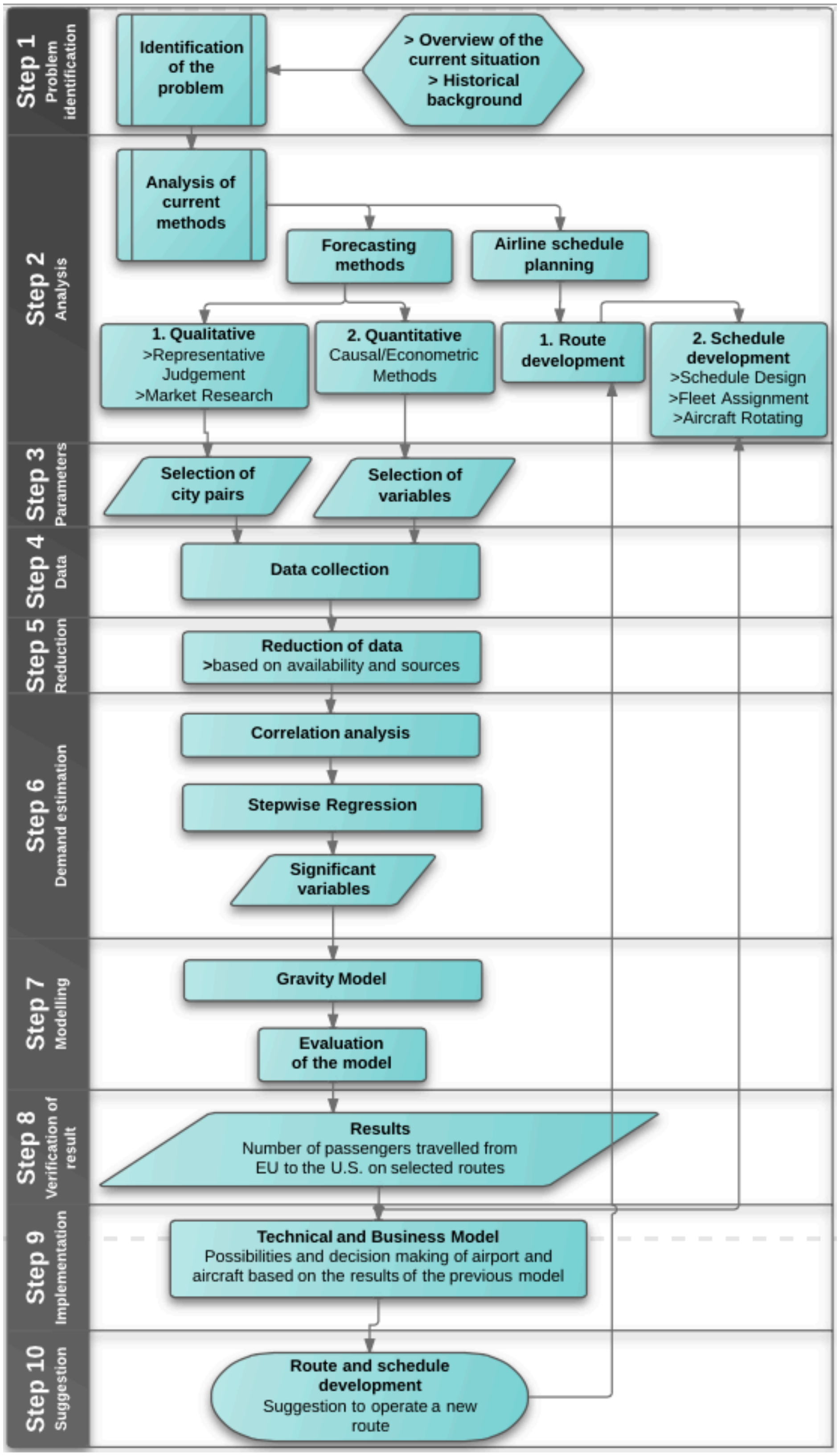


Figure 2.9: Sequence of steps to develop model for long-haul routes

Step 1 – Problem Identification

The sequence of steps began with a study of historical background and overview of the current situation of transatlantic air transportation. The objective and already known limitation of the problem were defined.

Step 2 – Analysis of Current Methods

Air travel forecasting techniques and airline schedule planning were reviewed. This thesis focuses on route development for airline schedule planning. Based on the analysis of techniques used to determine passenger demand, the following steps were set up.

Step 3 – Parameters

First, origin and destination airports of the long-haul routes network are defined. Then, explanation of parameters that may influence the demand presented. This definition is followed by collection of all needed data.

Step 4 – Data Collection

The collected data are split into dependent and independent variables.

Step 5 – Reduction of Data

Based on data available, the final data used to develop the model was chosen.

Step 6 – Demand Estimation

Correlation analysis was conducted to investigate statistical relationships between selected variables followed by stepwise regression with the significant variables in the regression model used as inputs in the gravity model.

Step 7 – Modeling

The gravity model was next developed and evaluated.

Step 8 – Verification of the Results

After successful completion of the gravity model, verification of the result has to be tested and the errors evaluated. Available data from Prague Airport will be used to compare with the forecasts.

Step 9 –Implementation

The utilization of Prague Airport and aircrafts were described.

Step 10 – Suggestions

As a result, the potential future route network was proposed. The limitations of the demand forecasting model documented. Suggestions were also made for future research.

Chapter 3: Airport Selection & Data Collection

This chapter focuses on description of airport selection and process of data collection that is one of the biggest challenges for later implementation of the forecasting model.

The aim of this chapter is to select airports from European and U.S. markets, and then the parameters for model development in the next chapter. For indication of the interaction and factors in air transportation that effects passengers' demand behavior and at the same time the airline planning, an influence list of all recognized impacts was created.

3.1 ROUTE DEVELOPMENT – PROPOSED APPROACH

Route development represents a very complex problem. There are some commercial applications, such as Sabre® AirVision™ that include software solution for route forecasting, planning and scheduling, fleet management, pricing and revenue management, and so on (Sabre 2013). These tools try to estimate passenger demand at a detailed level based on many parameters. This software is used in aviation industry mostly by big airlines.

The proposed approach follows the traditional planning process with application of iteration for more accurate results that is divided into sub-problems, where the each sub-problem is solved in sequence.

3.2 EXTERNAL IMPACTS ON PASSENGER DEMAND

The demand of air traffic is affected by a various factors. The following list tries to defined them:

- Location of the airport – distance, catchment area, modes of transport available
- Tourist attractiveness
- Marketing – alliance member, advertisements, deals, brand, packages, reward program, number of luggage
- Services – catering, cleanness, available business seats, first class, lounge,
- Technology – operating fleet, methods of check-in

Airline management can influence most of these factors. On the other hand, there are some factors that airlines cannot influence, such as:

- Increasing prices - as response to increasing price of fuel or airport fees
- Political situation – decision making of government, economy and the stability of the country

3.3 PERFORMANCE INDICATORS USED IN AIR TRANSPORT

For a better understanding of the following steps of the thesis, the basic measurements used in data collection are described in this section based on American Airlines (2013):

- **Revenue Passenger Miles (RPM)** is an indicator of production:

$$RPM = pax * miles \tag{3.1}$$

One RPM is defined as one paying passenger flying one mile. For example, 250 passengers flying 3,450 miles from Amsterdam to Boston generates 862,500 RPM

- **Available Seat Miles (ASM)** is a measure of supply or capacity:

$$ASM = seats * miles \tag{3.2}$$

One ASM measures one seat flying one mile. For example, on the flight Amsterdam- Boston flown by A330-300 with capacity 300 seats on distance 3,450 miles makes 1,035,000 ASM.

- **Load Factor (LF)** is a ratio of productivity to capacity expressed in percent:

$$LF = \frac{RPM}{ASM} \tag{3.3}$$

It can be imagined that as demand versus supply of a specific aircraft on a flight. In this given example, LF would be 0.83, or 83 percent. High LF does not always bring profit to an airline. This is obvious

from the example of the route between Prague and New York in 2009-2012. Other important factors that contribute to profit are occupancy by business passengers, operated aircraft, and operation of the route (e.g. aircraft routing).

- **Yield** is calculated as Revenue per Passenger Mile:

$$yield = \frac{revenue}{total\ RPMs} \tag{3.4}$$

Output of yield is expressed in currency per mile (or kilometer[2]).

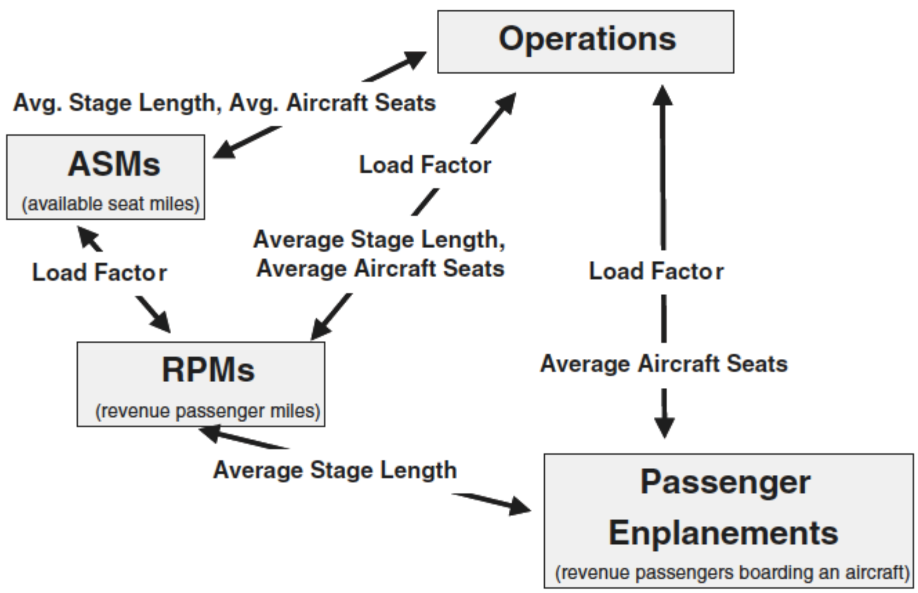


Figure 3.1: Relationships among the common passenger indicators

Source: (TRB 2007)

The interaction of performance indicators in air transport is presented in Figure 3.1.

2 based on inputs, e.g. in case of RPK, the yield is in $ per kilometer

3.4 AIRPORTS SELECTION

The process of selecting airports for future development of demand forecasting methodology was conducted by qualitative techniques. Because the thesis focuses on development of transatlantic route network, the selection is limited to the major airports in Europe and U.S. Hence, data are collected from the selected airports and will not cover all potential airports from Europe and U.S.

First, the judgmental method was used to investigate the six airports in U.S: New York (JFK), Boston (BOS), Chicago (ORD), Miami (MIA), Los Angeles (LAX), and San Francisco (SFO), followed by a few airports in Europe that have competing flights with Prague and passengers from Prague usually take connecting flights from Prague via these cities to U.S.: Vienna (VIE), Frankfurt (FRA), Paris (CDG), London (LHR), Amsterdam (AMS), Copenhagen (CPH), and Zurich (ZHR).

The Marketing Manager of Prague Airport endorsed this selection.

Then, the selection was supported by market analysis from Prague Airport and worldwide statistics.

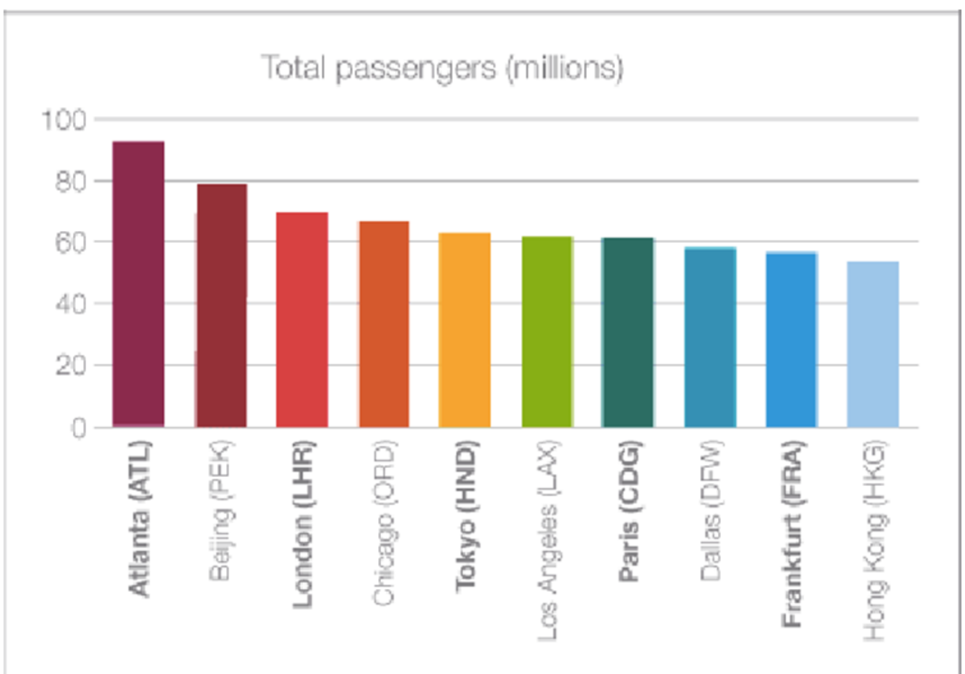


Figure 3.2: The busiest airport by passenger in terms of number of passengers in 2011

source: (Boeing 2012b)

Figure 3.2 shows that Atlanta (ATL), Chicago O'Hare (ORD), Los Angeles (LAX), and Dallas (DFW) airports are the four the biggest airports in U.S. in terms of total passengers per year. London Heathrow (LHR), Paris Charles De Gaulle (CDG), and Frankfurt (FRA) are the biggest European airports in terms of number of passengers per year.

Prague Airport (PRG) published potential route map and Marketing Information Data Transfer (MIDT) [3] based on the data from worldwide airline booking database - Global Distribution Systems (GDS).

Table 3.1 and Figure 3.3 describe potential market possibilities from Prague Airport to U.S airports, (Los Angeles, Chicago, San Francisco, Miami and Boston). They are evaluated as potential routes according the worldwide booking system by Prague (PRG 2012b).

Table 3.1: Potential market flights from Prague Airport to U.S. in 2012

adapted from (PRG 2012b).

DESTINATION	CODE	EXISTING ROUTE POTENCIAL (Passengers/Year)
Los Angeles	LAX	23 800
Chicago	ORD	19 600
San Francisco	SFO	19 600
Miami	MIA	17 800
Boston	BOS	16 500

[3] *"MIDT is a database that by accessing information from the Global Distribution Systems (GDS) captures booking transaction data from Passenger Name Records (PNR) to provide detailed information about the worldwide booking activities of airlines and travel agencies. MIDT was specifically designed to provide airlines with valuable competitive information so that they are able to make well informed decisions regarding existing and new route opportunities" (Market Analyser 2013)*

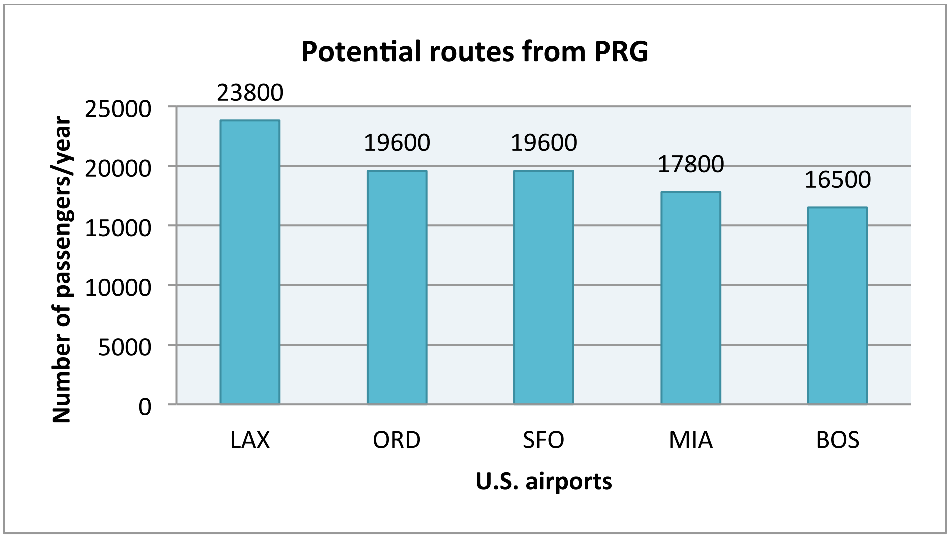


Figure 3.3: Potential market flights from PRG to U.S. in 2012

adapted from (PRG 2012b)

3.4.1 Competitors of Prague Airport

Figure 3.4 demonstrates the top 10 transfer points that were used during the travel from Prague Airport to U.S. in 2011, in terms of number of passengers. Among these airports are Frankfurt (FRA), London (LHR), Amsterdam (AMS), Paris (CDG), Munich (MUC), and Zurich (ZHR).

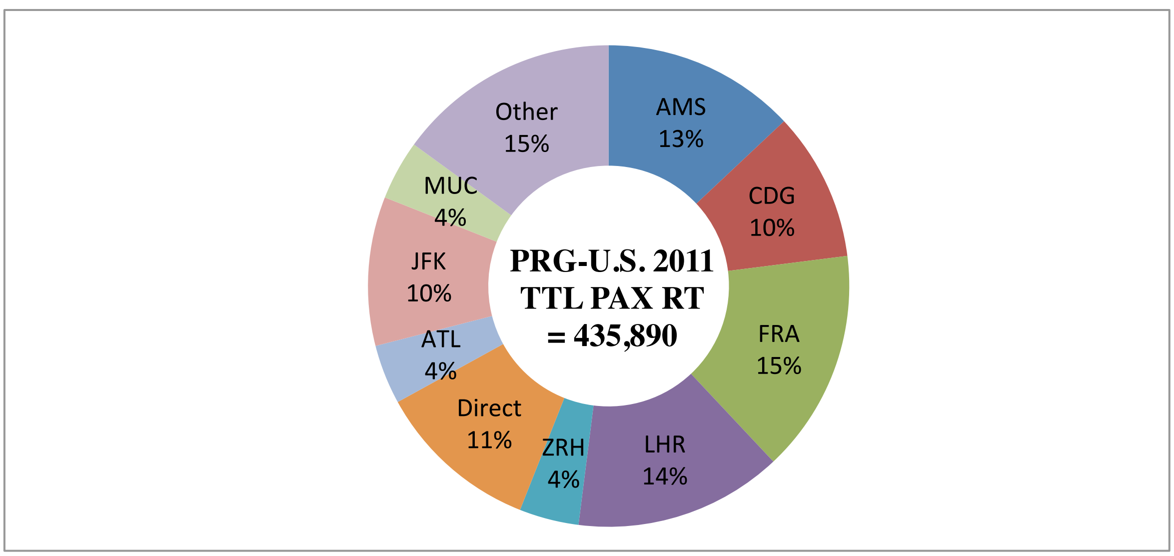


Figure 3.4: Top Transfer points PRG-North America to U.S. in 2011

source: (Prague Airport meeting 2012)

3.4.2 Summary

In summary, from Figure 3.1, Atlanta, Chicago O'Hare, Los Angeles, and Dallas were suggested. In addition, San Francisco, Miami and Boston were suggested from Table 3.1. New York (JFK) is not mentioned probably because of the direct flights that were operated from Prague Airport till November 2012 by Delta Airlines. New York also has 10 percent share in Figure 3.4. For the thesis Atlanta and Dallas are excluded, because most passengers use them as hubs, not final destinations.

After completing a comparison of the most used origins and destinations between Europe and U.S. for development of the model, the six airports were chosen: New York, Boston, Chicago, Miami, Los Angeles, and San Francisco (see in Table 3.2). The selected European airports are: Vienna, Frankfurt, Paris, London, Amsterdam, Copenhagen, Zurich, and Budapest (see Table 3.3).

Table 3.2: Final selection of U.S. airports

U.S. Airports	IATA code
Logan International Airport	BOS
John F. Kennedy International Airport	JFK
Los Angeles International Airport	LAX
Miami International Airport	MIA
Chicago O'Hare International Airport	ORD
San Francisco International Airport	SFO

Table 3.3: Final selection of European Airports

European Airports	IATA code
Amsterdam Airport Schiphol	AMS
Budapest Liszt Ferenc Inter. Airport	BUS
Charles De Gaulle Airport	CDG
Frankfurt am Main Airport	FRA
Heathrow Airport	LHR
Vaclav Havel Airport Prague	PRG
Vienna International Airport	VIE
Zurich Airport	ZRH

3.5 Selection of Variables

Identification of variables is an important step for causal modeling. In general, there are two types of variables: dependent and independent.

Dependent variables are typically the number of passengers (passenger trips) on an airport-pair route during a set period.

Independent variables are related mainly to two factors (Srinidhi 2009):

- Geo-economical factors describing economic activity,
- Geographical factors describing location impacts.

3.5.1 Methodology

Figure 3.5 describes the process of the selection of the variables for modeling. First, the historical data of the following variables was chosen based on the previous literature review. Variables were selected based on literature review and then reduced by data available.

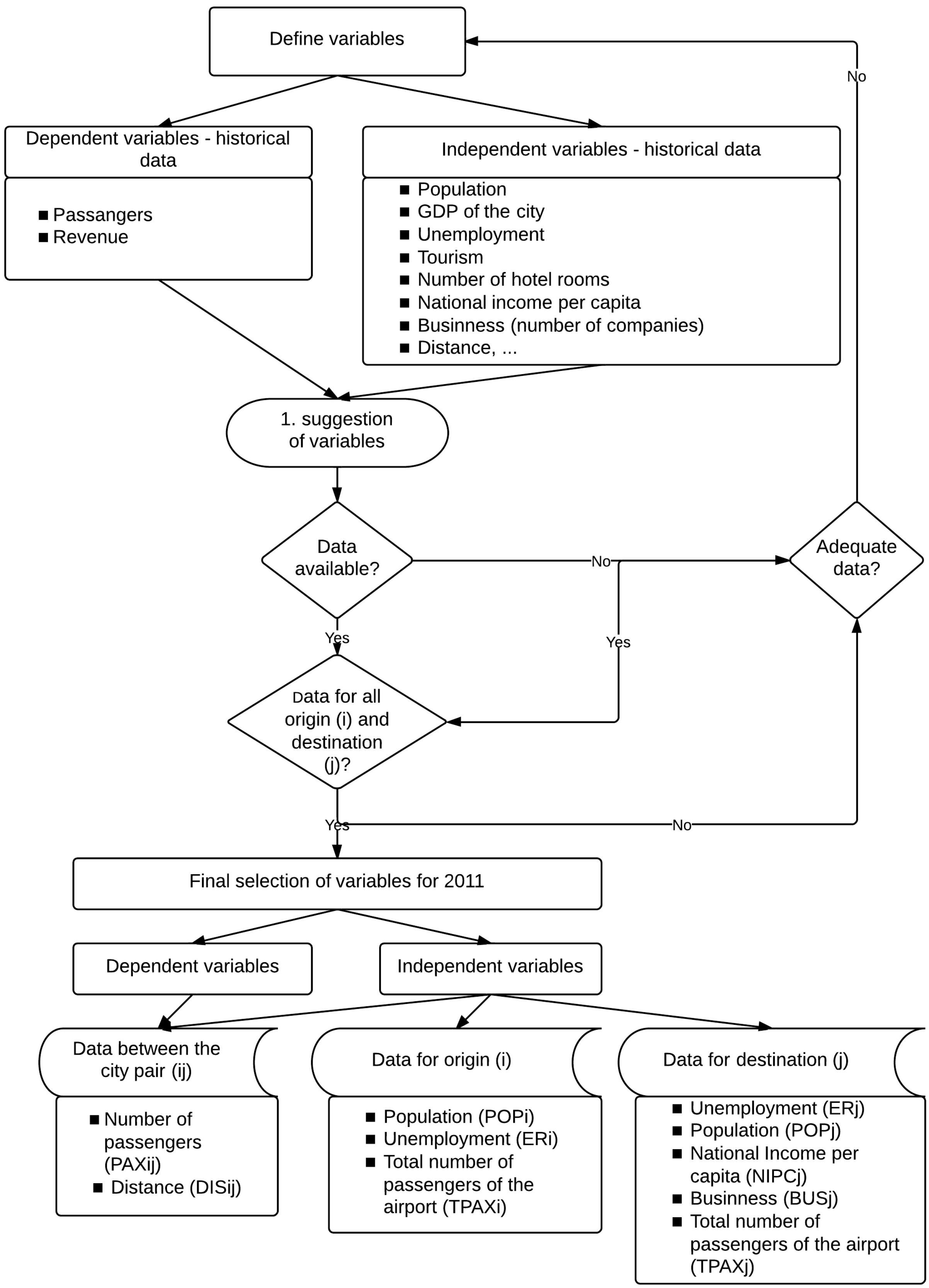


Figure 3.5: Process of selection variables for modeling

3.5.2 Summary

Based on previous research (Chang, 2011) ten variables are statistically significant in determining passenger flows between airport pairs. The final selection of variables is shown in Table 3.4.

Table 3.4: Defined variables

Dependent variables	Code	Type	Description
Number of passengers	PAX	Continuous	Number, one-way from European to U.S. airports
Independent variables	**Code**	**Type**	**Description**
Population	POP	Continuous	Number, in 2011
Unemployment rate	UER	Continuous	%, in 2011
Distance	DIS	Continuous	Distance between two selected cities, in air miles
Total passengers of the airport	TPAX	Continuous	Total passengers handled in 2011
National income	NIPC	Continuous	US$ (per capita) of the destination city in 2011
Business	BUS	Continuous	Total number of companies in the destination city in 2007

3.6 DATA SET

In general, data regarding airline planning and route development are not available to the public, because its represents highly classified and sensitive information of the airlines. Therefore, for the modeling work in this thesis, the data were obtained from different sources. Some real data were obtained from staff of the Prague Airport; some were obtained from airlines websites and air transport association websites. In addition, some data had to be collected according to the available fragmented information and then, estimated by indexing.

The focus of this study is on direct routes from European airports to U.S. airports. Hence, the data are limited to the airports identified in Tables 3.2 and 3.3.

Three types of data have been collected (See Figure 3.6):

- Demand data (obtained from Prague Airport's management)
- Supply data (airlines schedule)
- Geo-economical data (available statistics and databases)

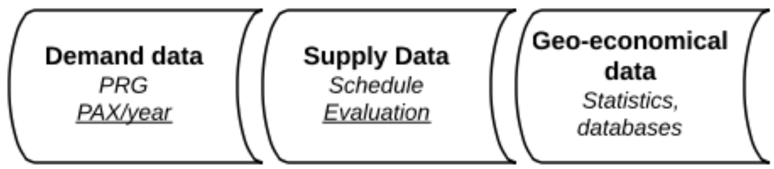


Figure 3.6: Types of data

3.6.1 Demand Data

In general, demand data signifies complete information of number of passengers based on the manifested airline schedules. One of the passenger demand sources is MIDT (Grosche 2009). According the Grosche (2009), the absolute passenger demand is estimated using gravity model.

The passenger demand data were obtained, in terms of total number of passengers travelling from Prague Airport to the various U.S. airports in 2011. The passenger demand data from the others European airports are obtained according to the method described in the next section.

3.6.2 Supply Data

Supply data were collected to fill up the missing demand data that were not possible to get, especially for trips originating from European airports other than Prague Airport. In short, the aim of the collection of supply data is to eventually estimate the total number of passengers travelling between the selected airport pairs.

The supply data depends on the flight schedules information available, such as the types of aircraft for specific flights. For the purpose of this thesis, the supply data was collected from Expedia (2013), one of the world's largest databases of airline's schedule. It is important to mention that only supply data of the direct flights were recorded. Next, these data were collected in two distinct periods:

airlines' summer timetable (August 2012) and winter timetable (March 2013). For each flight between a defined airport pair, the following data were collected on daily basis (for one week in the defined period):

- Date and departure time
- Airlines offering the flight
- Operated airline
- Type of aircraft

3.6.3 Geo-economic Data

Geo-economic data describes the demographic and economic activities of the city that serves the selected airports. The data used for the selected airports are shown in Tables 3.5 and 3.6.

Table 3.5: Geo-economic data of the origin airports

IATA code	City	Population 2011 POP	Unemployment rate 2011 UER (%)	Total passengers 2011 TPAX
VIE	Vienna	1731236	9.2	21106292
CDG	Paris	11800000	8.6	60970551
FRA	Frankfurt	691518	5.2	56 436 255
AMS	Amsterdam	780559	8.2	49754 910
ZRH	Zurich	390082	3.8	24337954

Table 3.6: Geo-economic data of the destination airports

IATA code	City	Population POP 2011	Unemployment rate UER 2011 (%)	Total PAX 2011	National income per capita in the city 2011 ($)	Total number of companies 2007
LAX	Los Angeles	3819702	11.4	61862052	28222	450108
SFO	San Francisco	805340	9.4	40800352	46777	105030
JFK	New York City	82444910	8.5	47854283	31417	944129
ORD	Chicago	2707120	13.6	66701241	27940	255502
BOS	Boston	625087	6.6	28907938	33158	49667
MIA	Miami	408750	10.5	38314389	20732	85146

3.7 METHODOLOGY OF ESTIMATION OF PASSENGER DEMAND FROM SUPPLY DATA

Figure 3.7 illustrates the process of estimation of passenger demand from the supply data. The goal was to estimate the total number of passengers between a defined airport pairs (one-way from Europe to U.S. airports) in 2011. Detailed description of the estimation procedure is followed. For easier symbolization, the number of passengers is mark as PAX in Figure 3.7.

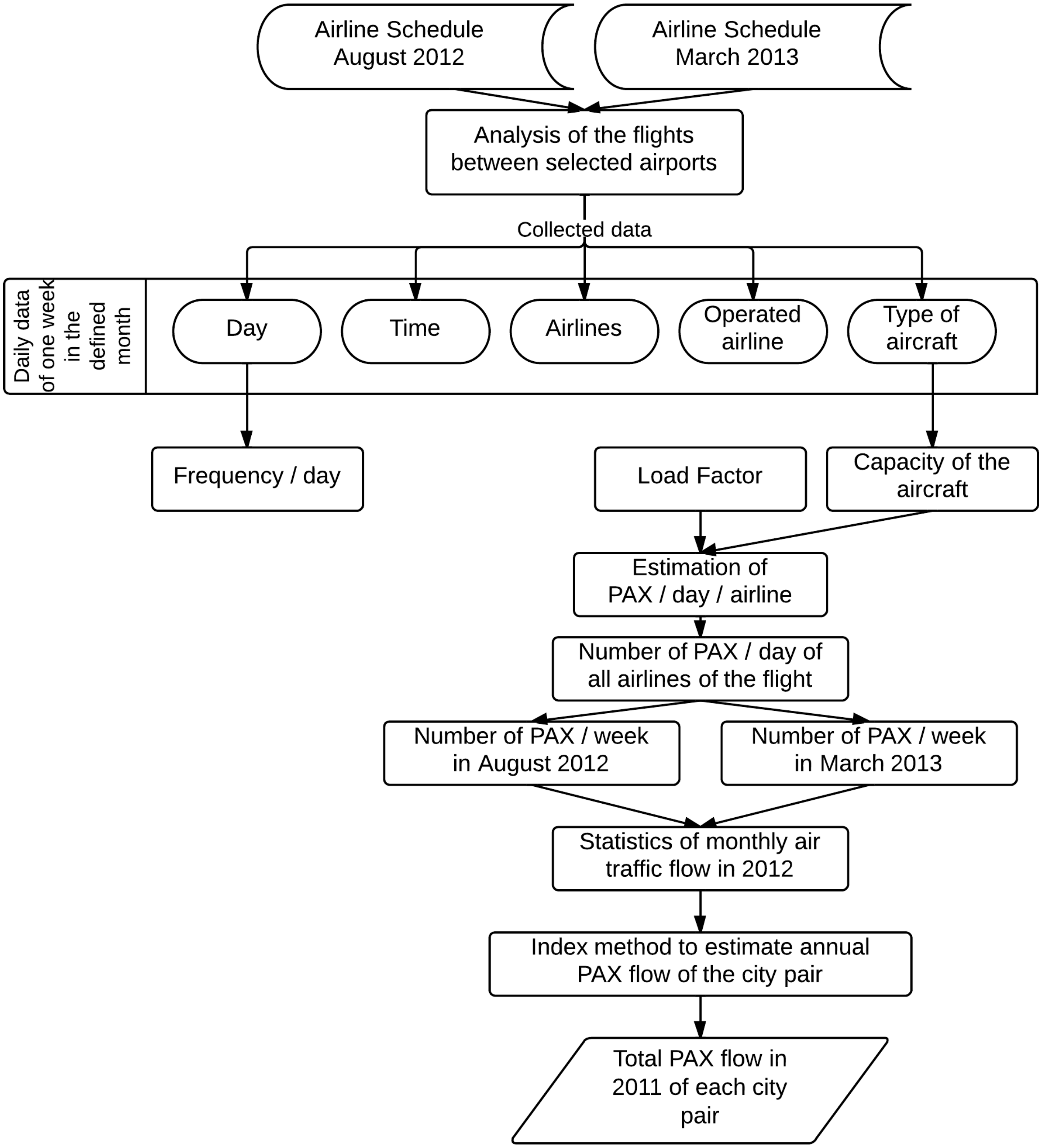


Figure 3.7: Process of estimation of passenger demand from the supply data

The total annual number of passengers in 2011 between each defined airport pair was calculated as follow:

1. Number of passengers of the unique flight *f* between airports *i* and *j*:

$$pax_{ij}^{f} = \max capacity * LF \tag{3.5}$$

where LF relevant for August 2012, March 2013 were used respectively. In case data for these months were not available, the annual LF in 2011 was taken into account. The capacity of aircraft (max number of seats) was obtained from the fleet information on the websites of the airlines.

2. Number of passengers a day:

$$pax_{ij}^{day} = \sum_{f=1}^{n} pax_{ij}^{f} \tag{3.6}$$

where *n* is the number of flights in the defined day

3. Number of passengers a week:

$$pax_{ij}^{week} = \sum_{day=1}^{7} pax_{ij}^{day} \tag{3.7}$$

4. Number of passengers a month:

$$pax_{ij}^{month} = pax_{ij}^{week} * 4.3, \tag{3.8}$$

where 4.3 is a constant that indicates the average number of the weeks in one month.

3.7.1 Index Method for Estimation of Annual Passenger Flow

The index method was calculated based on monthly Bureau of Transportation Statistics T-100 Market data that consisted of international data from the year 2012 of all airports and all carriers. All numbers are for scheduled services (US DOT 2013). Then, the estimation of supply data is calculated by

index method of monthly international passenger demand. The index method is based on set up the highest number of passengers travelled in the year as 1 (in other words 100%) and the other indexes are recalculated based on the international passenger demand of other months (created from percentage). Afterwards, the collected data of passenger demand from August 2012 and March 2013 are filled up in the table (see highlighted values in Table 3.7). The rest of the passenger demand is calculated based on this collected data and the index of the other months, according to equation (3.9)

$$pax_{ij}^{month\ x} = pax_{ij}^{month\ August,March} * x^{(August/March)} \quad (3.9)$$

Table 3.7: Sample of index method for estimating number of passengers between city-pairs

Month	International PAX demand (PAX/month)	Index x	FRA-MIA (PAX/ month)	FRA-JFK (PAX/ month)	FRA-BOS (PAX/ month)	FRA-ORD (PAX/ month)	FRA-LAX (PAX/ month)	FRA-SFO (PAX/ month)
1-2012	13,199,421	0,69	6,367	31,052	6,205	18,997	7,341	15,427
2-2012	11,729,831	0,75	5,658	27,595	5,514	16,882	6,524	13,709
3-2012	14,399,994	0,78	6,946	33,877	6,770	20,725	8,009	16,830
4-2012	14,018,004	0,79	6,762	32,978	6,590	20,176	7,796	16,384
5-2012	14,312,722	0,81	6,904	33,672	6,729	20,600	7,960	16,728
6-2012	15,605,186	0,82	7,528	36,712	7,336	22,460	8,679	13,709
7-2012	16,990,503	0,83	8,196	39,971	7,988	24,454	9,449	15,427
8-2012	16,636,022	0,84	8,025	39,137	7,821	23,944	9,252	14,911
9-2012	13,811,125	0,85	6,662	32,492	6,493	19,878	7,681	16,142
10-2012	13,343,965	0,92	6,437	31,393	6,273	19,205	7,421	15,596
11-2012	12,758,049	0,98	6,154	30,014	5,998	18,362	7,096	14,911
12-2012	13,929,226	1,00	6,719	32,769	6,548	20,048	7,747	16,280
Number of passengers:	***170,734,048***		***82,361***	***401,663***	***80,266***	***245,730***	***94,956***	***186,056***

Figure 3.8 shows the indexes of the international air transport from and to U.S. The statistics reflect highest passenger demand in summer months such as July and August. On the contrary, the lowest passenger demand is during February.

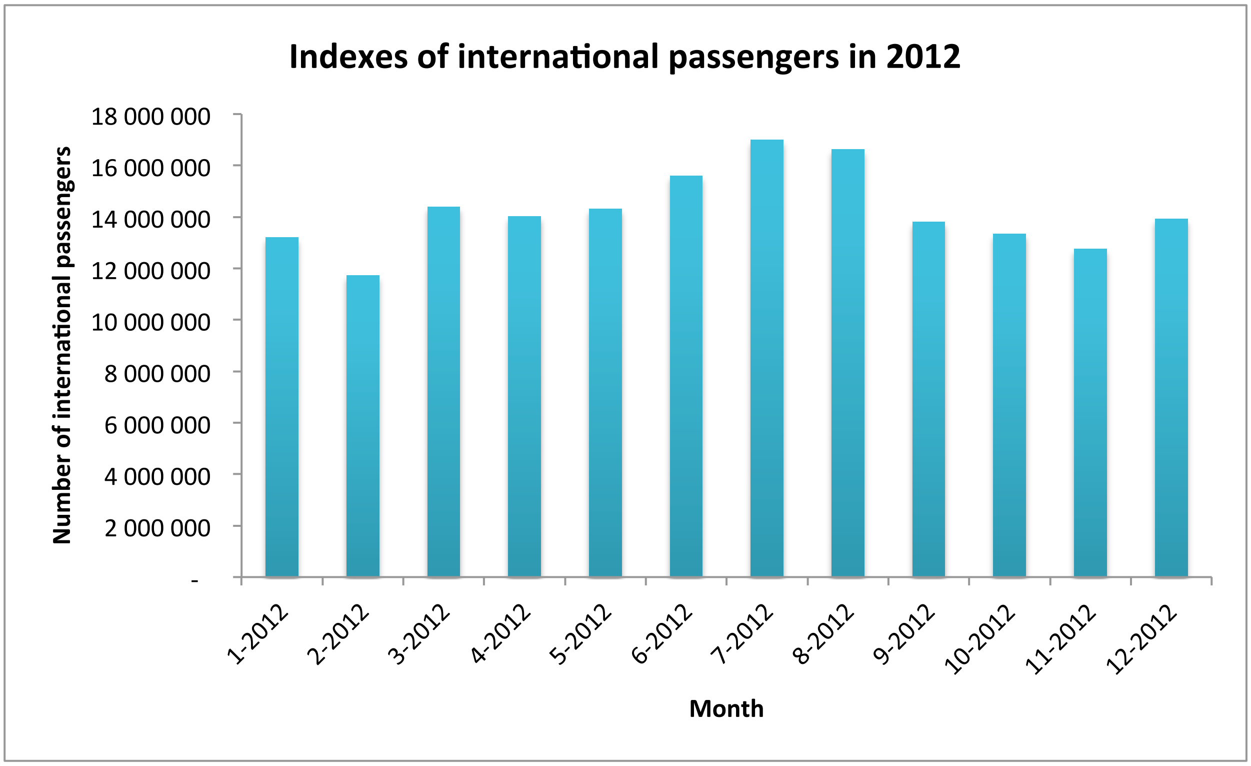


Figure 3.8: Number of international passengers in 2012 for calculation of the indexes

3.8 Limitation of Data

Data was mainly collected from available sources on the Internet, such as US DOT, U.S. Department of Commerce, Airports Council International or ACI, Boeing, ICAO, IATA and so on. The rest had to be estimated based on the statistics, predictions and by indexing.

The sample of the passengers demand data from Prague Airport to U.S. through the biggest hubs in Europe and U.S. was obtained from Prague Airport. This is not the actual demand for direct flights from Prague Airport to U.S.

3.9 SUMMARY

This chapter contains the description of the selection of the most important and significant airports for transatlantic air transportation from Europe to U.S. For instance, the biggest hubs in the Europe providing direct flights to the most attractive destination in U.S were included. However, the big U.S. hubs as Atlanta or Dallas that are used generally as hubs but not final destinations were not included. On the other side, Los Angeles International Airport, John F. Kennedy International Airport, Miami International Airport, Chicago O'Hare International Airport, San Francisco International Airport, Logan International Airport belongs to the biggest airports' hub and also as a final destination.

The biggest challenge of this chapter was the collection and estimation of the annual passenger demand from the detected European airports to the selected U.S. airports. All existing scheduled flights between defined airport-pairs had to be collected from available flight databases including date, time, operated airlines, co-shared airlines, operated aircraft and investigated their capacity based on configuration of the seats. For each airports-pair the extra sheet in Excel was created and based on the load factors, used operated aircraft and frequencies passenger demand was estimated. After completing database of data needed, for example index methods was used to core database of input data for further simulation. Additionally, the needed statistics data such as population, unemployment rate, total annual passengers of the origin and destination airports, national income, number of companies had to be collected for future modeling and demand estimation.

Chapter 4: Development of the Demand Forecasting Model

This chapter focuses on creation of a model for demand forecasting of transatlantic air transportation. First, a correlation analysis is presented in Section 4.1. Second, a regression analysis is performed to uncover significant variables. The main idea of regression is to find the relationships among the variables and use these relationships for making predictions.

Then, the gravity model is applied. In the last section of this chapter, the results of the gravity model are compared with the data provided by Prague Airport.

4.1 DEMAND FORECASTING MODELING

The models are applied to eight European airports and six airports in U.S. Supply data was collected and estimated for year 2012 and geo-economical and demand data for 2011. Detailed description of this process has been described in Chapter 3.

4.2 DATA INPUT

Data of 31 airport-pairs (origin and destination airports) were used as input for the correlation analysis and regression analysis (Table 4.1). These 31 airport-pairs are connected by direct flights. Airport pairs without a direct flight are not included in the table and in the correlation and regression analysis. As input, the following variables are used:

- Number of passengers, PAX_{ij} (passengers/year), 2011
- Distance between the origin and destination, DIS_{ij} (mi)
- Population of origin, POP_i (population), 2011
- Unemployment rate of origin, UER_i (%), 2011
- Total passengers of the origin airport, $TPAX_i$ (passengers/year), 2011
- Population of the destination, POP_j (population), 2011
- Unemployment rate of the destination, UER_j (%), 2011
- Total passengers of the destination airport, $TPAX_j$ (passengers/year), 2011
- National income of the destination, $NIPC_j$ (US dollars), 2011

- Business of the destination, BUS_j (number of companies), 2007

Table 4.1: Data set for correlation and regression analysis

i	j	PAX_{ij} (PAX/ year)	DIS_{ij} (mi)	POP_i (popula-tion)	UER_i (%)	$TPAX_i$ (PAX/ year)	POP_j (popula-tion)	UER_j (%)	$TPAX_j$ (PAX/ year)	$NIPC_j$ (US dollars)	BUS_j (no. of companies)
AMS	BOS	91409	3450	780559	8,2	49754910	625087	6,6	28907938	33158	49667
AMS	JFK	308579	3630	780559	8,2	49754910	8244910	8,5	47854283	31417	944129
AMS	LAX	125323	5560	780559	8,2	49754910	3819702	11,4	61862052	28222	450108
AMS	ORD	111278	4110	780559	8,2	49754910	2707120	13,6	66701241	27940	255502
AMS	SFO	61475	5460	780559	8,2	49754910	805340	9,4	40800352	46777	105030
CDG	BOS	74590	3440	11800000	8,6	60970551	625087	6,6	28907938	33158	49667
CDG	JFK	581162	3620	11800000	8,6	60970551	8244910	8,5	47854283	31417	944129
CDG	LAX	278612	5650	11800000	8,6	60970551	3819702	11,4	61862052	28222	450108
CDG	MIA	141432	4580	11800000	8,6	60970551	408750	10,5	38314389	20732	85146
CDG	ORD	154181	4140	11800000	8,6	60970551	2707120	13,6	66701241	27940	255502
CDG	SFO	183158	5570	11800000	8,6	60970551	805340	9,4	40800352	46777	105030
CPH	ORD	69867	4260	1213822	7,9	22725517	2707120	13,6	66701241	27940	255502
FRA	BOS	82326	3660	691518	5,2	56436255	625087	6,6	28907938	33158	49667
FRA	JFK	406864	3840	691518	5,2	56436255	8244910	8,5	47854283	31417	944129
FRA	LAX	97392	5790	691518	5,2	56436255	3819702	11,4	61862052	28222	450108
FRA	MIA	83428	4820	691518	5,2	56436255	408750	10,5	38314389	20732	85146
FRA	ORD	252037	4330	691518	5,2	56436255	2707120	13,6	66701241	27940	255502
FRA	SFO	204670	5680	691518	5,2	56436255	805340	9,4	40800352	46777	105030
LHR	BOS	425381	3250	14900000	8,3	69433565	625087	6,6	28907938	33158	49667
LHR	JFK	1538893	3440	14900000	8,3	69433565	8244910	8,5	47854283	31417	944129
LHR	LAX	1076444	5440	14900000	8,3	69433565	3819702	11,4	61862052	28222	450108
LHR	MIA	250336	4410	14900000	8,3	69433565	408750	10,5	38314389	20732	85146
LHR	ORD	534749	3940	14900000	8,3	69433565	2707120	13,6	66701241	27940	255502
LHR	SFO	258562	5350	14900000	8,3	69433565	805340	9,4	40800352	46777	105030
VIE	JFK	60534	4220	1731236	9,2	21106292	8244910	8,5	47854283	31417	944129
ZRH	BOS	70499	3730	390082	3,8	24337954	625087	6,6	28907938	33158	49667
ZRH	JFK	209163	3920	390082	3,8	24337954	8244910	8,5	47854283	31417	944129
ZRH	LAX	56068	5920	390082	3,8	24337954	3819702	11,4	61862052	28222	450108
ZRH	MIA	110771	4870	390082	3,8	24337954	408750	10,5	38314389	20732	85146
ZRH	ORD	70491	4430	390082	3,8	24337954	2707120	13,6	66701241	27940	255502
ZRH	SFO	65413	5820	390082	3,8	24337954	805340	9,4	40800352	46777	105030

Source: ACI (2012), CDG (2012), Gemeente Amsterdam (2012), LAWA (2012), Stadt Zürich (2012), U.S. Bureau of Labor Statistics (2012), U.S. Census Bureau (2012), WebFlyer (2013), Wien.at (2012), Zurich Airport (2012)

4.3 Correlation Analysis

Correlation analysis is a statistical process that identifies the degree of relationship between variables:

- Correlation *r* can be positive (from 0 to 1) or negative from (from -1 to 0).
- If $|r|=0$: no correlation
- If $|r|\leq 0.4$: weak correlation
- If $|r|=0.5\text{-}0.8$: medium correlation
- If $|r|\geq 0.9$: high correlation

In this thesis, correlation analysis was modeled in Excel by the excel function CORREL:

The equation for the correlation coefficient is:

$$Correl\ (X,Y) = \frac{\sum_{i=1}^{n} \sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum_{i=1}^{n} \sum(x-\bar{x})^2 \sum_{i=1}^{n} \sum(y-\bar{y})^2}} \tag{4.1}$$

where *x*, *y* are variables.

Table 4.2: Correlation analysis of variables

Variables	PAX_{ij} (PAX/ year)	DISij (mi)	POP_i (popu-lation)	UER_i (%)	$TPAX_i$ (PAX/ year)	POP_j (popu-lation)	UER_j (%)	$TPAX_j$ (PAX/ year)	$NIPC_j$ (US $)	BUS_j (no. of companies
POPI	1	-0.643	0.720	-0.059	0.017	-0.050	-0.025	-0.052	0.153	0.565
ERI	-0.643	1	-0.580	-0.091	-0.003	-0.040	-0.023	-0.086	-0.201	0.295
TPAXI	0.720	-0.580	1	-0.108	0.046	-0.079	0.022	-0.100	0.130	0.494
POPJ	-0.059	-0.091	-0.108	1	0.030	0.364	-0.129	0.996	0.254	0.441
ERJ	0.017	-0.003	0.046	0.030	1	-0.878	0.364	0.059	0.449	-0.054
TPAXJ	-0.050	-0.040	-0.079	0.364	-0.878	1	-0.294	0.326	-0.326	0.126
NIPCJ	-0.025	-0.023	0.022	-0.129	0.364	-0.294	1	-0.127	-0.205	-0.054
BUSJ	-0.052	-0.086	-0.100	0.996	0.059	0.326	-0.127	1	0.212	0.445
DISij	0.153	-0.201	0.130	0.254	0.449	-0.326	-0.205	0.212	1	-0.217
PAXij	0.565	0.295	0.494	0.441	-0.054	0.126	-0.054	0.445	-0.217	1

First, 31 values of eight independent variables (population of origin, population of destination, unemployment of origin, unemployment of destination, total passengers of the origin airport, total passengers of the destination airport, national income of destination and business of destination) were analyzed by the CORREL function in Excel. Additionally, distance between airport pairs was evaluated.

From Table 4.3 it is obvious that strongly related pairs of variables are business of destination BUS_j and population of destination POP_j.

The other strong related variables are total passengers of the airport of origin $TPAX_i$ and population of origin POP_i. The last strongly related is unemployment of origin UER_i and population of origin.

4.4 Regression Analysis

Using the data of 31 airport pairs as 31 observations, stepwise regression modeling (linear model, forward selection) was conducted to relate the dependent variable PAX_{ij} with other independent variables. The estimated coefficients describe the positive or negative influence of the independent variables on PAX_{ij}.

4.4.1 Data Input

The linear regression was modeled in NLogit 4.0

.

```
--> REGRESS;           Lhs = PAXij;
                       Rhs =
POPi,ERi,TPAXi,POPj,ERj,TPAXj,NIPCj,BUSj,distd;ALG=stepwise; list...
+---------------------------------------------------------------------+
| Stepwise Regression                                                 |
| Dependent variable     = PAXIJ                                      |
| Number of observations =       31                                   |
| Number of regressors   =        9                                   |
| Degrees of freedom     =       21                                   |
| Predictor variables are:                                            |
| POPI     ERI      TPAXI    POPJ     ERJ      TPAXJ    NIPCJ    BUSJ |
| DISTD                                                               |
+---------------------------------------------------------------------+
```

First, the file of inputs data had to be created and then the variables of right side and left side of the regression equation has to be defined in the NLogit program.

4.4.3 Description of Stepwise Regression

The stepwise regression method selects the best combination of independent variables that best predicts the dependent variable.

4.4.4 Steps of Stepwise Regression

Step 1

```
+-----------------------------------------------------+
| Step  1 of Stepwise Regression                      |
| Ordinary    least squares regression                |
| Model was estimated Apr 02, 2013 at 08:32:18AM      |
| LHS=PAXIJ    Mean                 =   259196.4      |
|              Standard deviation   =   318877.7      |
| WTS=none     Number of observs.   =         31      |
| Model size   Parameters           =          2      |
|              Degrees of freedom   =         29      |
| Residuals    Sum of squares       =   .2076463E+13  |
|              Standard error of e  =   267585.8      |
| Fit          R-squared            =   .3193017      |
|              Adjusted R-squared   =   .2958294      |
| Model test   F[  1,    29] (prob) =  13.60 (.0009)  |
| Diagnostic   Log likelihood       =  -430.3664      |
|              Restricted(b=0)      =  -436.3283      |
|              Chi-sq [   1]  (prob) =  11.92 (.0006) |
| Info criter. LogAmemiya Prd. Crt. =   25.05691      |
|              Akaike Info. Criter. =   25.05673      |
| Mallows Cp =    34.477                              |
+-----------------------------------------------------+
```

Results of the Step 1

```
+--------------------------------------------------------------------+
| Analysis of Variance for the Current Regression                    |
| Source       Deg.Fr.      Sum of squares      Mean Square       F  |
| Regression      1     974026393887.77200 ****************  13.60   |
```

```
| Residual       29    2076462774862.75900 ****************              |
| Total          30    3050489168750.53100 ****************              |
| Variable entered this step = POPI    , Deleted =                        |
+-------------------------------------------------------------------------+
```

Variable	Coefficient	Standard Error	t-ratio	P[|T|>t]	Mean of X
POPI	.02840342	.00770102	3.688	.0009	.559798D+07
Constant	100194.534	64561.8729	1.552	.1315	

Step 1: Variable POP_i was obtained with p value of 0.0009.

Step 2

```
+-----------------------------------------------------+
| Step  2 of Stepwise Regression                      |
| Ordinary    least squares regression                |
| Model was estimated Apr 02, 2013 at 08:32:19AM      |
| LHS=PAXIJ    Mean                 =   259196.4      |
|              Standard deviation   =   318877.7      |
| WTS=none     Number of observs.   =         31      |
| Model size   Parameters           =          3      |
|              Degrees of freedom   =         28      |
| Residuals    Sum of squares       =   .1386719E+13  |
|              Standard error of e  =   222543.7      |
| Fit          R-squared            =   .5454109      |
|              Adjusted R-squared   =   .5129402      |
| Model test   F[  2,    28] (prob) =  16.80 (.0000)  |
| Diagnostic   Log likelihood       =  -424.1087      |
|              Restricted(b=0)      =  -436.3283      |
|              Chi-sq [  2]  (prob) =  24.44 (.0000)  |
| Info criter. LogAmemiya Prd. Crt. =   24.71813      |
|              Akaike Info. Criter. =   24.71752      |
| Mallows Cp =    19.371                              |
+-----------------------------------------------------+
```

Results of Step 2

```
+-------------------------------------------------------------------+
| Analysis of Variance for the Current Regression                   |
| Source       Deg.Fr.       Sum of squares      Mean Square      F |
| Regression      2    1663770007599.12300 ****************  16.80 |
| Residual       28    1386719161151.40800 ****************         |
| Total          30    3050489168750.53100 ****************         |
| Variable entered this step = BUSJ    , Deleted =                  |
+-------------------------------------------------------------------+
```

Variable	Coefficient	Standard Error	t-ratio	P[|T|>t]	Mean of X
POPI	.02965633	.00641352	4.624	.0001	.559798D+07
BUSJ	.46127316	.12360316	3.732	.0009	340722.419
Constant	-63985.3444	55224.3587	-1.159	.2564	

Step 2: Variable BUS_j was obtained with p value of 0.0009.

Step 3

```
+----------------------------------------------------+
| Step  3 of Stepwise Regression                     |
| Ordinary    least squares regression               |
| Model was estimated Apr 02, 2013 at 08:32:20AM     |
| LHS=PAXIJ    Mean                 =   259196.4     |
|              Standard deviation   =   318877.7     |
| WTS=none     Number of observs.   =         31     |
| Model size   Parameters           =          3     |
|              Degrees of freedom   =         28     |
| Residuals    Sum of squares       =   .1386719E+13 |
|              Standard error of e  =   222543.7     |
| Fit          R-squared            =   .5454109     |
|              Adjusted R-squared   =   .5129402     |
| Model test   F[  2,    28] (prob) =  16.80 (.0000) |
| Diagnostic   Log likelihood       =  -424.1087     |
|              Restricted(b=0)      =  -436.3283     |
|              Chi-sq [  2]  (prob) =  24.44 (.0000) |
| Info criter. LogAmemiya Prd. Crt. =   24.71813     |
|              Akaike Info. Criter. =   24.71752     |
| Mallows Cp =    19.371                             |
+----------------------------------------------------+
```

Results of Step 3

```
+--------------------------------------------------------------------+
| Analysis of Variance for the Current Regression                    |
| Source        Deg.Fr.       Sum of squares      Mean Square      F |
| Regression       2    1663770007599.12300 ****************  16.80 |
| Residual        28    1386719161151.40800 ****************         |
| Total           30    3050489168750.53100 ****************         |
| Variable entered this step =          , Deleted =                  |
| **********> This is the final equation <**********                 |
+--------------------------------------------------------------------+
```

Variable	Coefficient	Standard Error	t-ratio	P[|T|>t]	Mean of X
POPI	.02965633	.00641352	4.624	.0001	.559798D+07
BUSJ	.46127316	.12360316	3.732	.0009	340722.419
Constant	-63985.3444	55224.3587	-1.159	.2564	

Stepwise regression model contains, as a result, the independent variables POP_i, BUS_j with p values of 0.001 and 0.009 respectively. These two variables are significant variables and will be used as input for gravity model in the other section.

4.5 GRAVITY MODEL

For modeling of demand the gravity model was chosen. Gravity model calculates the number of passengers between two airports in relation to the productivity and attractiveness of these airport-pairs.

The purpose of the gravity modeling is to predict the PAX_{ij} of transatlantic flights between defined airports in Europe and U.S.

4.5.1 Data Input

Before the modeling, it is necessary to adjust the data set. Based on the output of stepwise regression, for gravity modeling POP_i and BUS_j are used as the variables that describe the trip generation rates of the origin i and the attractiveness of destination j respectively. Then, the distance (DIS_{ij}) is used as the impedance of travel between i and j. Because only 31 airport pairs have direct flights, the following operations were made to create a 6x6 matrix for gravity model calculation:

1) The airport pairs with only one direct flight in time period of collecting the data were removed. Two European airports are neglected: Budapest and Vienna. Copenhagen offers just one direct flight too in the time period of data collection, but according the flight schedule in March 2013, there is a new direct flight to San Francisco (Expedia 2013). It demonstrates the increasing potential of this airport in the question of transatlantic flights. Copenhagen is therefore included in the gravity model development. As Table 4.3 illustrates, the airport pairs with indirect flights are assigned values of PAX_{ij}=0.001 trip/year that represents no passengers in the stated route.

Table 4.3: Review of airports with only one direct flight to the selected destinations in 2011

PAX_{ij} (passengers/year)	BOS	JFK	LAX	MIA	ORD	SFO
BUD	0.001	0.001	0.001	0.001	0.001	0.001
CPH	0.001	0.001	0.001	0.001	69867	0.001
VIE	0.001	60534	0.001	0.001	0.001	0.001

2) In case of missing direct flights, the value 0.001 was put into the matrix of estimated number of passengers (PAX_{ij}) (see Table 4.4)

Table 4.4: The matrix of estimated number of passengers (PAX_{ij}) as gravity model input

PAX_{ij} (passengers/year)	BOS	JFK	LAX	MIA	ORD	SFO
AMS	89122	300857	122187	0.001	102908	59937
CDG	72724	573733	275051	139624	150323	180817
CPH	0.001	0.001	0.001	0.001	69867	0.001
FRA	80266	401663	94956	82361	245730	204670
LHR	414737	1500386	357313	247136	521369	255257
ZRH	65191	203929	54665	100697	68727	63776

3) For the airport pairs without adirect flight, the number 999999 miles was added into the distance matrix *(DIS_{ij})*. This number was set up to force the city-pairs with indirect connection to have PAX_{ij}=0 in the final table. (see Table 4.5)

Table 4.5: The distance matrix *(DIS_{ij})* as gravity model input

	BUS_i (no. of companies)		POP_i (population)	DIS_{ij} (mi)	BOS	JFK	LAX	MIA	ORD	SFO
BOS	49667	AMS	780559	AMS	3450	3630	5560	999999	4110	5460
JFK	944129	CDG	11800000	CDG	3440	3620	5650	4580	4140	5570
LAX	450108	CPH	1213822	CPH	999999	999999	999999	999999	4260	999999
MIA	85146	FRA	691518	FRA	3660	3840	5790	4820	4330	5680
ORD	255502	LHR	14900000	LHR	3250	3440	5440	4410	3940	5350
SFO	105030	ZRH	390082	ZRH	3730	3920	5920	4870	4430	5820

4.5.2 Calculation

The number of passenger-trips between two airports T_{ij} is calculated by following gravity model equation according to Cheu (2012):

$$T_{ij} = a_i b_j \frac{POP_i BUS_j}{DIS_{ij}^x} \tag{4.8}$$

where:

a_i	is an airport specific trip production constant (to be calibrated)
b_j	is an airport specific trip attraction constant (to be calibrated)
POP_i	is population of origin airport
BUS_j	is business of destination airport
DIS_{ij}	is destination between the airport of origin and destination
x	is calibration parameter (exponential coefficient)

Production constant a_i and attraction constant b_j is calibrated as:

1) Assume all $a_i = 1$,
2) Calculate all b_j using $b_j = \left[\sum_{i=1}^{n} \frac{a_i POP_i}{DIS_{ij}^x} \right]^{-1}$ (4.9)
3) Calculate all a_i using $a_i = \left[\sum_{j=1}^{n} \frac{b_j BUS_j}{DIS_{ij}^x} \right]^{-1}$ (4.10)
4) Calculate $T_{ij} = a_i b_j \frac{POP_i BUS_j}{DIS_{ij}^x}$ for all i and j
5) Repeat steps (2)-(4) until all a_i, b_j and T_{ij} converges

The result of the application of a gravity model is a demand matrix T_{ij}.

4.5.3 Model Development in Matlab

The gravity model with variables POP_i and BUS_j was tested in Matlab first. Model was tested with x=1 and for different number of iterations k=5, k=50, *and* k=2000. The results showed that after the second iteration T_{ij} do not change significantly. The matrix of average absolute relative errors does not change significantly either.

The gravity model is also tested in Excel running for the first 10 iterations to see how the gravity models behave.

4.5.4 Model Development in Excel

4.5.4.1 Modification of Data Input

The model requests modification of data to balance the trips from all origins and all destinations. The sum of the number of passenger-trips in a row should equal to the number of passenger-trips emanating from that airport; the sum of the number of passenger-trips in a column corresponds to the total trips attracted to that airport (Bruno & Improta 2008).

These conditions can be written as:

$$\sum_i PAX_{ij} = \sum_{i=1}^{6} POP_{i\_pax} \tag{4.11}$$

$$\sum_j PAX_{ij} = \sum_{j=1}^{6} BUS_{j\_pax} \tag{4.12}$$

In Table 4.6, passenger trips are summed and this sum (7099979 passenger trips/year) is used to try to keep model into balance.

Table 4.6: The matrix of estimated number of passengers (PAX_{ij}) and summarization of passengers trips

The matrix of estimated number of passengers (PAX_{ij}) as gravity model input							
PAX_{ij}	BOS	JFK	LAX	MIA	ORD	SFO	**O_i_pax**
AMS	89122	300857	122187	0.001	102908	59937	675012
CDG	72724	573733	275051	139624	150323	180817	1392271
CPH	0.001	0.001	0.001	0.001	69867	0.001	69867
FRA	80266	401663	94956	82361	245730	204670	1109646
LHR	414737	1500386	357313	247136	521369	255257	3296198
ZRH	65191	203929	54665	100697	68727	63776	556985
D_j_pax	722040	2980569	904172	569818	1158923	764457	**7099979**

In Table 4.7 below, the sum of trips of origins and destinations are balanced by utilization of coefficient 3.7574.

Table 4.7: Modification of data by utilization of coefficient

Modification of data								
T_{ij}	BOS	JFK	LAX	MIA	ORD	SFO	POP_i	
AMS	676	705	991	63096	778	977	780559	
CDG	65	703	1004	849	783	992	11800000	
CPH	63096	63096	63096	63096	801	63096	1213822	
FRA	709	737	1024	884	811	1008	691518	
LHR	645	675	974	823	752	961	14900000	
ZRH	720	749	1042	891	826	1028	390082	$\sum POP_i$
BUS_j	49667	944129	450108	85146	255502	105030		**29775981**
						$\sum BUS_j$	**1889582**	

1st iteration (k=1)								
T_{ij}	BOS	JFK	LAX	MIA	ORD	SFO	a_i	POP_i
AMS	1324	25235	12262	32	6618	2867	1	48338
CDG	20065	382336	183007	35603	99469	42651	1	763132
CPH	22	438	299	49	10001	69	1	10879
FRA	1119	21373	10517	2003	5624	2461	1	43096
LHR	26515	502887	238194	46338	130676	55621	1	1000230
ZRH	622	11859	5828	1121	3115	1361	1	23906
bj	0.00002	0.00002	0.00003	0.00003	0.0000š	0.00003		**1889582**
BUSj	49667	944129	450108	85146	255502	105030	**1889582**	**3.7574**

2nd iteration (k=2)								
T_{ij}	BOS	JFK	LAX	MIA	ORD	SFO	a_i	POP_i
AMS	5270	100416	48722	126	21161	11391	16.1478	187086
CDG	76468	1456824	696292	135773	304545	162287	15.4626	2832190
CPH	607	12049	8222	1355	220948	1895	111.5768	245076
FRA	4425	84511	41523	7926	17868	9716	16.0459	165969
LHR	97349	1846016	873091	170240	385446	203888	14.8966	3576031
ZRH	2500	47686	23400	4509	10064	5466	16.3175	93626
bj	0.000002	0.000002	0.000002	0.000002	0.000001	0.000002		**7099979**
BUSj	186620	3547502	1691251	319930	960032	394643	**7099979**	

4.5.5 Modeling

Two variables of destination were tested, POP_j and BUS_j, which represent the attractions at U.S. airports. For each variable, x was varied and all the coefficients a_i and b_j were calibrated. The accuracy of the gravity model was judged based on the absolute average percentage error between the prediction (T_{ij}) and observed data (PAX_{ij}).

Different values of calibration parameter x were tested. After testing various values of x, the range of x=0.8 to 1.2 has been found with the smallest values of the error. Out of this range the results were not acceptable.

Based on this observation, the gravity model was tested with three x values: 0.8, 1.0, and 1.2.

4.5.5.1 Evaluation of Errors of Simulation with Variables of Population and Business

Table 4.8 shows the results of 10 iterations of gravity model calibration with three different values of x for variables POP_i and BUS_j. From the results it is obvious that the smallest average absolute relative error of all airport-pairs in the error matrix is obtained in configuration with $x=0.8$. In Table 4.8, it is also shown that the smallest value of error is in the second iteration and the error magnitude is 0.79.

Table 4.8: Error results of with different values of x for POP_i and BUS_j

k	*x=0.8*	*x=1.00*	*x=1.2*
1	0.80	0.80	0.82
2	0.77	0.78	0.79
3	0.79	0.79	0.81
4	0.79	0.80	0.81
5	0.79	0.80	0.81
6	0.79	0.80	0.81
7	0.79	0.80	0.81
8	0.79	0.80	0.81
9	0.79	0.80	0.81
10	0.79	0.80	0.81
Average	0.79	0.80	0.81

Figure 4.1 plots the curves of the average absolute percentage errors of the matrixes. It can be observed that the smallest errors are placed in the second iteration.

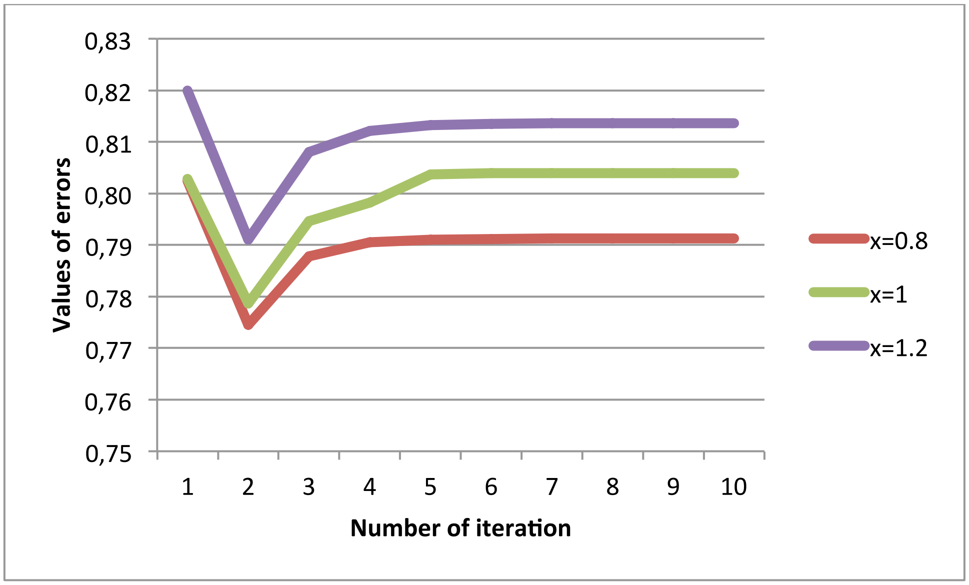


Figure 4.1: Graph of average absolute percentages errors for POP_i and BUS_j

4.5.5.2 Evaluation of Errors of Simulation with Variables of Population of Origin and Destination

Because of previous results with massive placing of errors, POP_j was used to substitute BUS_j. The variable of population of destination for the second gravity model was chosen based on correlation analysis that showed big correlation between BUS_j and POP_j.

Table 4.9: Average absolute percentages errors results of simulation of 10 iterations with different values of x for POP_i and POP_j

k	*x=0.8*	*x=1*	*x=1.2*
1	1.2460	1.2458	1.2455
2	0.8130	0.8166	0.8177
3	0.8246	0.8301	0.8319
4	0.8266	0.8327	0.8348
5	0.8270	0.8332	0.8354
6	0.8270	0.8333	0.8355
7	0.8270	0.8334	0.8356
8	0.8270	0.8334	0.8356
9	0.8270	0.8334	0.8356
10	0.8270	0.8334	0.8356
Average	0.8672	0.8725	0.8743

Table 4.9 depicts errors of the second gravity model. As is obvious from the table, the errors of the second model are even bigger than in first.

Figure 4.2 illustrates the close behavior of the errors in all defined x values. After the second iteration the errors are almost the same.

Figure 4.2: Graph of error results of simulation of 10 iterations for POP_i and POP_j

4.5.5.3 Modeling based on Variables of Population and Business

As a result from previous section, the model based on POP_i and BUS_j is run with $x=0.8$. The results are displayed in Table 4.10 for iterations $k=2,3,$ and 4.

Table 4.10: Modeling two variables: POP_i and BUS_j with $x=0.8$, $k=2,3,4$

2nd iteration (k=2)								
T_{ij}	BOS	JFK	LAX	MIA	ORD	SFO	a_i	POP_i
AMS	7395	97766	46108	67	25922	9740	1.8012	186998
CDG	109621	1449139	673220	74209	381155	141775	1.7622	2829118
CPH	733	10091	6694	624	232834	1394	10.7065	252369
FRA	6346	84090	40159	4334	22369	8491	1.8292	165788
LHR	139500	1835568	843834	93012	482220	178049	1.6970	3572183
ZRH	3586	47457	22635	2466	12602	4777	1.8605	93524
bj	0.000013	0.000014	0.000020	0.000017	0.000012	0.000020		**7099979**
POPj	267180	3524111	1632650	174711	1157101	344225	**7099979**	

3rd iteration (k=3)								
T_{ij}	BOS	JFK	LAX	MIA	ORD	SFO	a_i	POP_i
AMS	7397	97795	46112	67	25165	9741	3.2136	186278
CDG	109569	1448433	672765	74166	369745	141680	3.1415	2816358
CPH	845	11631	7714	719	260456	1606	22.0104	282971
FRA	6343	84053	40134	4331	21700	8485	3.2612	165047
LHR	139441	1834765	843305	92963	467809	177939	3.0255	3556222
ZRH	3584	47434	22620	2465	12225	4774	3.3168	93102
bj	0.000008	0.000008	0.000011	0.000010	0.000007	0.000011		**7099979**
POPj	267180	3524111	1632650	174711	1157101	344225	**7099979**	

4th iteration (k=4)								
T_{ij}	BOS	JFK	LAX	MIA	ORD	SFO	a_i	POP_i
AMS	7397	97797	46112	67	25034	9741	5.756	186149
CDG	109560	1448308	672684	74159	367772	141663	5.626	2814147
CPH	865	11907	7896	736	265235	1644	40.356	288283
FRA	6343	84046	40129	4331	21585	8484	5.840	164918
LHR	139431	1834622	843211	92954	465316	177919	5.418	3553452
ZRH	3584	47430	22618	2464	12159	4774	5.940	93029
bj	0.000004	0.000004	0.000006	0.000005	0.000004	0.000006		**7099979**
POPj	267180	3524111	1632650	174711	1157101	344225	**7099979**	

4.5.5.4 Error Matrix

The error matrix expresses the average absolute percentage error between each airport pair. In the error matrix, in Table 4.11, it can be seen that the errors are quite big. The airport pairs with errors smaller than 0.30 are highlighted by green color. The most accurate results are in the cases of airport pairs involving Paris (CDG) and London (LHR) as the origins. As a destination, San Francisco (SFO) has two pairs with errors smaller than0.20. In the bottom of the each table (SUM of E) is the average absolute percentage error of all the airport pairs.

Table 4.11: Error matric of iteration *k=2,3,4* and *x=0.8*

Error E_2	BOS	JFK	LAX	MIA	ORD	SFO
AMS	0.94	0.67	0.27		0.79	0.81
CDG	0.05	1.54	1.53	0.03	1.03	0.10
CPH					2.16	
FRA	0.94	0.79	0.56	0.90	0.93	0.95
LHR	0.77	0.23	1.44	0.31	0.26	0.20
ZRH	0.96	0.77	0.57	0.96	0.85	0.91
SUM of E	0.77					

Error E_3	BOS	JFK	LAX	MIA	ORD	SFO
AMS	0.94	0.67	0.27		0.80	0.81
CDG	0.05	1.54	1.53	0.03	0.94	0.10
CPH					2.60	
FRA	0.94	0.79	0.56	0.90	0.93	0.95
LHR	0.77	0.23	1.44	0.31	0.29	0.20
ZRH	0.96	0.77	0.57	0.96	0.86	0.91
SUM of E	0.79					

Error E_4	BOS	JFK	LAX	MIA	ORD	SFO
AMS	0.94	0.67	0.27		0.80	0.81
CDG	0.05	1.54	1.53	0.03	0.92	0.10
CPH					2.69	
FRA	0.94	0.79	0.56	0.90	0.93	0.95
LHR	0.77	0.23	1.44	0.31	0.30	0.20
ZRH	0.96	0.77	0.57	0.96	0.86	0.91
SUM of E	0.79					

4.5.5.5 Summary of Result Simulation with Variables of Population and Business

Then, the modeling was made in Excel for k=(1,10). The errors are not reduced with higher number of iterations. The final values were changing till the third iteration and then they were almost stable. It might be said they are oscillating.

4.6 SUMMARY

The gravity model may not be not suitable for demand forecasting with the defined variables used in this chapter, because of large magnitude of errors. Nevertheless, the best gravity model developed in this thesis will be applied to forecast the transatlantic air passenger demands from Prague Airport in the next Chapter.

Chapter 5: Implementation – Prague Airport

Based on the evaluation of the average absolute percentage error matrixes in the previous chapter, the most suitable gravity model was selected to forecast air passenger demand between selected European airports and U.S. airports.

The results are compared with available data from Prague Airport.

In addition, the example of aircraft for these routes is recommended and a simple timetable is proposed, as an example of airline planning.

5.1 Gravity Model Application

From the previous chapter, the gravity with x=0.8, attraction variable BUS_j was identified as the most suitable for application for passenger demand forecasting between Prague Airport and U.S. airports.

1) Table 5.1 shows the inputs of the gravity model with generation and attraction variables, airport specific coefficients and distances.

Table 5.1: Modeling of configuration: variable BUS_j and POP_i, $x=0.8$, $k=2$

2nd iteration (k=2)								
T_{ij}	BOS	JFK	LAX	MIA	ORD	SFO	a_i	POP_i
AMS	5270	100416	48722	126	21161	11391	16.1478	187086
CDG	76468	1456824	696292	135773	304545	162287	15.4626	2832190
CPH	607	12049	8222	1355	220948	1895	111.5768	245076
FRA	4425	84511	41523	7926	17868	9716	16.0459	165969
LHR	97349	1846016	873091	170240	385446	203888	14.8966	3576031
ZRH	2500	47686	23400	4509	10064	5466	16.3175	93626
bj	0.0000015	0.0000016	0.0000023	0.0000020	0.0000014	0.0000022		**7099979**
BUSj	186620	3547502	1691251	319930	960032	394643	**7099979**	

2) Three value of a_i (Table 5.2, Table 5.3 and Table 5.4) are chosen from Table 5.1: Minimum a_i, average of a_i, and maximum a_i.

Table 5.2: Passenger demand of Prague Airport calculated with configuration: a_{min}, $x=0.8, k=2$

T_{prg} for a_{min}							
a_{min}	14.89656788						Input data
	BOS	JFK	LAX	MIA	ORD	SFO	POP_{prg}
T_{prg_j}	7,128.13	136,147.72	68,569.13	12,867.57	29,086.11	16,082.03	1257158
b_j	0.0000015	0.0000016	0.0000023	0.0000020	0.0000014	0.0000022	
BUS_j modified	186620	3547502	1691251	319930	960032	394643	

Table 5.3: Passenger demand of Prague Airport calculated with configuration: $a_{average}$, $x=0.8, k=2$

T_{prg} for $a_{average}$							
$a_{average}$	31.7412						Input data
	BOS	JFK	LAX	MIA	ORD	SFO	POP_{prg}
T_{prg_j}	15,188.43	290,099.78	146,105.21	27,417.86	61,975.88	34,267.14	1257158

Table 5.4: Passenger demand of Prague Airport calculated with configuration: a_{max}, $x=0.8, k=2$

T_{prg} for a_{max}							
a_{max}	111.5768483						Input data
	BOS	JFK	LAX	MIA	ORD	SFO	POP_{prg}
T_{prg_j}	53,390.46	1,019,760.63	513,589.97	96,379.44	217,858.02	120,456.09	1257158

3) Because the results based on previous selection of a_i are very different, the average absolute relative errors of using the minimum a_i=14.8966 is estimated and used for the route development. The average absolute relative errors of using the minimum a_i=14.8966 for all airport pairs are displayed in Table 5.5.

Table 5.5: Evaluation of the minimal value of error of a_i

Error E_2	BOS	JFK	LAX	MIA	ORD	SFO	a_i	Average E_{ri}
AMS	0.94	0.67	0.27		0.79	0.81	16.1478	0.70
CDG	0.05	1.54	1.53	0.03	1.03	0.10	15.4626	0.71
CPH					2.16		111.5768	2.16
FRA	0.94	0.79	0.56	0.90	0.93	0.95	16.0459	0.85
LHR	0.77	0.23	1.44	0.31	0.26	0.20	14.8966	0.54
ZRH	0.96	0.77	0.57	0.96	0.85	0.91	16.3175	0.84

Based on the previous steps, the forecasted passenger demand is evaluated according the smallest of $E_r = 0.54$ in error matrix (from LHR) in Table 5.5. As a result, the number of passengers in Table 5.6 is considered for route development and schedule planning. Figure 5.1 compares the results of the gravity model and the available MIDT data from 2011.

Table 5.6: Results of demand forecasting from Prague Airport

Estimated demand (PAX/year)	BOS	JFK	LAX	MIA	ORD	SFO
PRG	7124	136059	68505	12858	27875	16067

Figure 5.1: Comparison of the result of the gravity model and available MIDT data

Based on the result of gravity model forecasts, routes are suggested for the three destinations with the biggest potential. As is shown in Table 5.7, the biggest potential is to fly to JFK, LAX and ORD.

Table 5.7: Suggestion of the routes based on the results of gravity model

Destination	Number of passengers/year
JFK	136059
LAX	68505
ORD	27875

5.3 AIRCRAFT SUGGESTION

This section focuses on the challenge to suggest new routes and choose suitable aircraft. The aircraft is suggested for Czech Airlines fleet, because it is the one operating out of Prague Airport.

A330 belongs among the most fuel efficient aircraft. Because Czech Airlines is starting new long haul routes to the Asian market (June 2013) and they are going to operate four A330 aircrafts, the same aircraft is chosen for the routes to U.S.

As is mentioned in Table 5.8, there are two configurations of A330: A330-200 and A330-300.

Table 5.8: Configuration of A330-200 and A330-300

Specification	Airbus A330-200		Airbus A330-300	
Engines	2		2	
Thrust per engine	68,000-72,000 lb_f		68,000-72,000 lb_f	
MTOW	230 tonnes		230 tonnes	
Range	13,100 km	7,400 nm	11,000 km	6,500 nm
Cruise speed	M0.83		M0.83	
Typical seats	246 passengers		300 passengers	
Business class	30 lie-flat seats (60x20.25)		30 lie-flat seats (60x20.25)	
Economy Comfort Class	39 standard seats (35x17.5)		40 standard seats (35x17.5)	
Economy class	174 standard seats (31x17.5)		222 standard seats (31x17.5)	

For the route transatlantic routes is used configuration A330-300. The reasons are following:

- Is already in the fleet:
    - own experience (current planning routes of A330-300: 4 aircrafts – service to Moskva, Yekaterinburg, Soul, and Almata in Kazachstan),
    - can substitute in case some aircraft go away,
    - no need extra stuff for maintenance (lower maintenance cost, reduced cost for training for different aircraft[4]).
- Experience on the same routes of other airlines (Swiss for routes: ZHR-BOS, ZHR-ORD, ZHR-MIA, Delta for route AMS-JFK, SAS for CPH-ORD)

The route PRG-LAX might be operated by A330-200 after detailed research because of bigger range.

Table 5.9: Current numbers of operated, ordered, and delivered A330-300 in March 2013

Total orders	622
Total deliveries	456
In operation	449

Source: Airbus (2013)

[4] Each aircraft the certification is needed for crew, maintenance, handling, and so on

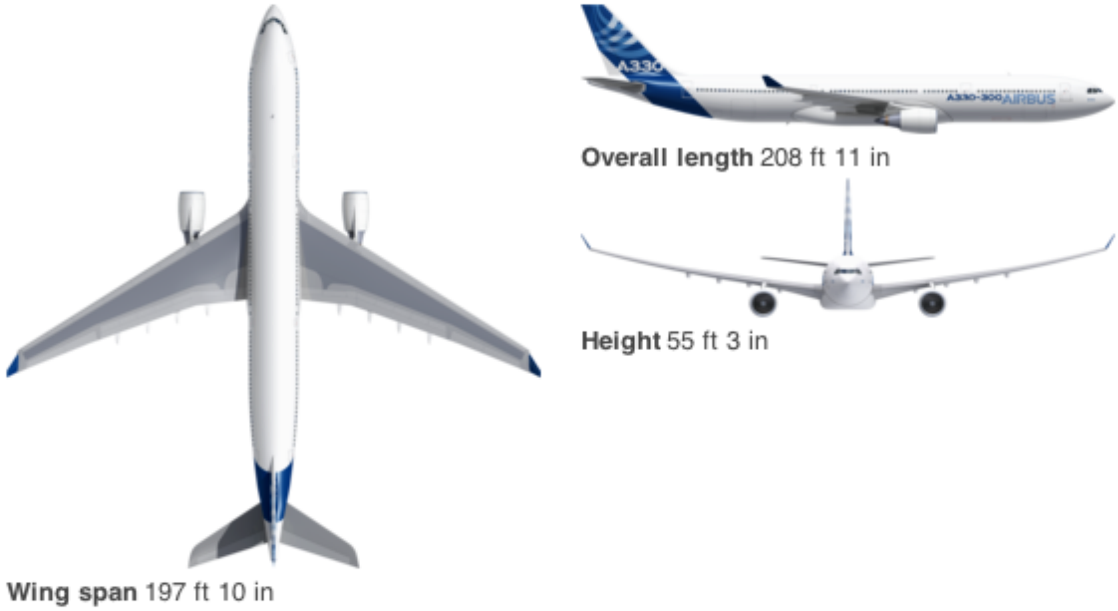


Figure 5.2: Dimensions of A330-300

Source: Airbus (2013)

5.4 Suggestion of Frequencies for Flights to New York, Los Angeles, and Chicago

Frequencies are suggested according to the LF of European biggest airlines, according Table 5.10.

Table 5.10: Examples of load factors of European airlines and their LF

Load factor on flight to U.S.	
Lufthansa	85.7%
Air France & KLM	80.9%
British Airways &Iberia	83.3%

As Table 5.11 demonstrates, the average *LF* is 83%. The number of passengers were calculated with *LF*=80%, 85%, 90% and based on the results the frequencies of flights is set up.

Table 5.11: Suggestion of frequency of flights

PRG	PAX/ year)	LF 80% (PAX/ year)	LF 85% (PAX/ year)	LF 90% (PAX/ year)	LF 80% (PAX/ week)	LF 85% (PAX/ week)	LF 90% (PAX/ week)	Frequency	Days
JFK	136059	566.91	533.56	503.92	10.87	10.23	9.66	11	1,1,2,2,3,4,4,5,6,7,7
LAX	68505	285.44	268.65	253.72	5.47	5.15	4.87	6	1,2,3,4,5,6
ORD	27875	116.15	109.31	103.24	2.23	2.10	1.98	3	1,3,5

5.5 Examples of Airline Schedule

In Tables 5.12 and 5.13, examples of arrival and departure timetable are created based on the flight frequencies proposed in Table 5.11. The timetable is just for illustration of how the transatlantic route network from PRG to JFK, LAX and ORD could look like. It can be improved by other planning and optimization methods, and incorporating all operational constraints, but this is not objective of this thesis.

Table 5.12: Example of arrival timetable of Prague Airport

ARRIVAL

Flight	**Departure time (hh:mm)**	**Travel time**	**Arrival time**	**Aircraft**	**Days**
PRG-JFK	13:00	8:30	15:30	A330-300	1,2,3,4,5,6,7
PRG-JFK	17:00	8:30	19:30	A330-300	1,2,4,7
PRG-LAX	10:00	13:20	14:20	A330-300/A330-200	1,2,3,4,5,6
PRG-ORD	11:00	10:00	14:00	A330-300	1,3,5

Table 5.13: Example of departure timetable of Prague Airport

DEPARTURE

Flight	**Departure time**	**Travel time**	**Arrival time**	**Aircraft**	**Days**
JFK-PRG	17:00	7:30	6:30	A330-300	1,2,3,4,5,6,7
JFK-PRG	21:00	7:30	10:30	A330-300	1,2,4,7

LAX-PRG	16:00	11:20	12:20	A330-300/A330-200	1,2,3,4,5,6
ORD-PRG	15:30	8:45	7:15	A330-300	1,3,5

The duration of the flight was calculated according the official timetable of the airlines (KLM, Swiss, etc.) that operate this flight from airports closest to the PRG and are using aircraft A330-300 or by comparing these results and using equation for speed, time and length. The duration has to be calculated for both direction because of head wind and tale wind that influence flight time about one hour and more.

5.6. SUMMARY

This chapter applied the gravity model to forecast air passenger demand from Prague Airport to selected U.S. airports. It shows how the gravity model may be utilized in airlines route planning. Based on the forecasts, routes from PRG to JFK, LAX and ORD are suggested, with weekly frequencies of 11, 6 and 3 flights, respectively. The recommended aircraft for these routes is A330. Timetables for these routes have also been proposed.

Chapter 6: Conclusion

This chapter summarizes the research presented, underlines its contributions, and gives recommendations for future research. Section 6.1 concludes the outcomes and objectives of the thesis. Section 6.2 summarizes the contributions. In contrary, Section 6.3 mentions possible limitations and proposes for the future. Last, Section 6.4 motivate for future research and improvements.

6.1 SUMMARY OF THE RESEARCH

This thesis reviews air travel forecasting techniques and airline schedule planning. Based on this review, this study introduces a correlation-regression-gravity model sequential methodology to estimate the passenger demand based on qualitative and quantitative methods of demand forecasting that are presented in the Chapter 2. Chapter 3 describes a process of selection of airports for transatlantic demand analysis and possible future operations for direct flights from Prague Airport. Then, the detailed procedure of data collection is explained including defining the variables. Chapter 4 shows the systematics process of developing the demand forecasting model based on correlation analysis, liner regression analysis and gravity model. Finally, in Chapter 5, the developed gravity model is applied to forecast demand for transatlantic flights out of Prague Airport. Three airports, namely JFK, LAX and ORG are recommended as U.S. destinations, with 11, 6 and 3 flights per week using A330.

6.2 CONTRIBUTION

Among the most significant contributions of this thesis are:

- The research provides information and justifies the importance of establishing Prague Airport as a hub for transatlantic flights.
- A methodology to determine the passengers demand based on qualitative and quantitative methods was established. Especially, the three causal-econometric methods was introduced in sequence, tested and evaluated.
- This research showed a process of how to find potential origins and destinations of an airlines network.

- This research demonstrates a methodology of performing demand forecast for transatlantic flights with the lack of historical commercial sensitive data, the possibilities of gathering publicly available data from various sources and the methods to process them.
- This research developed an approach and simple technique to help airlines to forecast the passenger demand in the defined market.
- This research showed examples of schedule development as a beginning of airline schedule planning process.

6.3 LIMITATIONS

The biggest challenge of this research was to determine which data are needed for the demand modeling and to collect the data needed.

The results of the gravity model may not be accurate enough for eventual decisions on route development and airline schedule planning. The errors of the gravity model can be due to:

- The estimated data are not precise enough, although the value was close to the historical demands of indirect flights from Prague Airport to the selected airports in U.S.
- The variables were not the best that represent the generation of the origins and attractiveness of the destinations. More variables should be explored if more data is available. The impedance function may not be the best as well.
- The gravity model is not the best solution for demand forecasting.

6.4 FUTURE RESEARCH

Based on the list of limitations, there are many possibilities on how to improve the gravity model, or to use other models. One of the options is to validate model by collecting more data or try to use more accurate data. Another option is to use different variables. Among the other challenges is to expand model to include route revenue as the dependent variable.

In the case of route development and airline schedule planning, there are a lot of options for modification and optimization. Nowadays, a lot of optimization and integrated tools exist and are applied to airline scheduling.

References

Abdelghany, A. and Abdelghany, K. (ed). (2009). *Modeling Applications in the Airline Industry*. Farnham, England: Ashgate Publishing, Ltd..

ACI - Airports Council International (2012). *ACI Releases its 2011 World Airport Traffic Report: Airport Passenger Traffic Remains Strong as Cargo Traffic Weakens - Aug 27, 2012* [online]. Available from: <http://www.aci.aero/News/Releases/Most-Recent/2012/08/27/ACI-Releases-its-2011-World-Airport-Traffic-Report-Airport-Passenger-Traffic-Remains-Strong-as-Cargo-Traffic-Weakens>. [Accessed 16 November 2012].

Airbus (2013). *A330-300* [online]. Available from: <http://www.airbus.com/aircraftfamilies/passengeraircraft/a330family/a330-300/>. [Accessed 20 April 2013].

Airliners.net (2008). *Czech Airlines Celebrating 85th Anniversary* [online]. Available from: <http://www.airliners.net/aviation-forums/general_aviation/print.main?id=4170279>. [Accessed 16 April 2013].

AirlineUpdate.com (2013). *Czech Airlines* [online]. Available from: <http://www.airlineupdate.com/content_public/airlines/airline_profiles/airlines_czech/csaczechairlines.htm>. [Accessed 16 April 2013].

American Airlines (2013). *Basic measurements in the Airline Business* [online]. Available from: <http://www.aa.com/i18n/amrcorp/corporateInformation/facts/measurements.jsp>. [Accessed 14 March 2013].

Belobaba, P. (2006). *Airline Management*. (PDF). Available from: <http://www.core.org.cn/NR/rdonlyres/Aeronautics-and-Astronautics/16-75JSpring-2006/398AF266-A71E-4A65-ACB6-2D55A27CF505/0/lect5.pdf>. [Accessed 20 October 2012].

Belobaba, P., Odoni, A. and Barnhart, C. (eds). (2009). *The global airline industry*. Wiltshire, United Kingdom: John Wiley.

Boeing (2012). *Aviation Policy and Geopolitics 2012*. Boeing Commercial Airplanes: Seattle. Available from: < http://www.boeing.com/assets/pdf/commercial/pdf/geopolitical_trends.pdf>. [Accessed: 24 March 2013].

Boeing (2012b). *Infrastructure* [online]. Available from: <http://www.boeing.com/assets/images/commercial/cmo/images/cmo_infrastructure_chart2_lrg.gif>. [Accessed 14 April 2013].

Boeing (2013). *Current Market Outlook 2012-2031* [online]. Available from: <http://active.boeing.com/commercial/forecast_data/index.cfm>. [Accessed 30 January 2013].

CAPA - Centre for Aviation (2012). *State aid: Commission approves restructuring aid for Czech Airlines* [online]. Available from: <http://centreforaviation.com/members/direct-news/state-aid-commission-approves-restructuring-aid-for-czech-airlines-83256>. [Accessed 16 April 2013].

Carryer, S. (2007). *Air Transport in the 1930s*. Pelgrane Press Limited [online]. October, p.-. Available from: <http://www.pelgranepress.com/seepagexx/october2007-simonc.html>. [Accessed 7 October 2012].

CDG - Aéroport de Paris Charles de Gaulle (2012). Statistiques annuelles [online]. Available from: <http://www.aeroport.fr/les-aeroports-de-l-uaf/stats-paris-charles-de-gaulle.php>. [Accessed 16 November 2012].

Chang, L. (2012). *International air passenger flows between pairs of APEC countries: A non-parametric regression tree approach*. Journal of Air Transport Management [online]. 20, p.4–6. Available from: <http://www.sciencedirect.com/science/article/pii/S0969699711000561>. [Accessed 28 November 2012].

Cheu, R. (2012). *Trip Distribution*, lecture notes distributed in the topic CE 6306 Infrastructure Engineering. University of Texas at El Paso, El Paso.

Czech Airlines (2013). *History* [online]. Available from: <http://usa.czechairlines.com/en/portal/quicklinks/about-us/history.htm>. [Accessed 22 October 2012].

Daniels, P. (2013). *Charles Lindbergh, The Famous Aviator in History*. Avaliable from: < http://history1900s.about.com/od/people/a/Charles-Lindbergh.htm >. [Accessed: 20 March 2013].

EAFAS - European Academy For Aviation Safety (1999). *ETOPS History*. (Power Point). Available from: <http://www.coscapsa.org/Operations/Courses/ETHIST98.PPT.>. [Accessed 7 October 2012].

Eurofound (2009). *Czech Airlines cuts wages in preference to dismissals* [online]. Available from: <http://www.eurofound.europa.eu/eiro/2009/05/articles/cz0905029i.htm>. [Accessed 16 April 2013].

Expedia, Inc. (2013). *Expedia, Inc*. [online]. Available from: <http://www.expediainc.com>. [Accessed 16 April 2013].

FAA - Federal Aviation Administration (2012). *Press Release – Airline Passenger Travel to Nearly Double in Two Decades* [online]. Available from: <http://www.faa.gov/news/press_releases/news_story.cfm?newsId=13394>. [Accessed 13 April 2013].

Francis, G., Dennis, N., Ison, S. (2007). *The transferability of the low-cost model to long-haul airline operations*. Tourism Management [online]. 28, p.391-398. Available from: <http://www.sciencedirect.com/science/article/pii/S0261517706000690>. [Accessed 6 December 2012].

Gemeente Amsterdam (2012). Bureaua Onderzoek Statistiek [online]. Available from: <http://www.os.amsterdam.nl/>. [Accessed 16 November 2012].

Greene, W. (2007). LIMDEP : Version 9.0 / Econometric modeling guide. Vol.1. ed. Plainview, NY: Econometric Software.

Grosche, T. (2009). *Computational Intelligence in Integrated Airline Scheduling*. 1st. ed. Berlin, Germany: Springer.

Grosche, T., Rothlauf, F. and Heinzl, A. (2007). *Gravity models for airlines passenger volume estimation*. *Journal of Air Transportation Management* [online]. 13, p.175-183. Available from: <http://www.sciencedirect.com/science/article/pii/S0969699707000178>. [Accessed 26 February 2013].

Hansen M. and Wei, W. (2006). *An Aggregate demand model for air passenger traffic in the hub-and-spoke network*. Transportation Research Part A: Policy and Practice [online]. 40, p.841-851. Available from: <http://www.sciencedirect.com/science/article/pii/S0965856406000218>. [Accessed 26 April 2012].

IATA - International Air Transport Association (2012). *Fact Sheet: IATA - International Air Transport Association* [online]. Available from: <http://www.iata.org/pressroom/facts_figures/fact_sheets/pages/iata.aspx>. [Accessed 6 March 2013].

ICAO - International Civil Aviation Organization (2011). *Doc 9889, Airport Air Quality Manual*. ICAO: Montréal, Quebec, Canada.

ICAO - International Civil Aviation Organization (2013a). *Forecasts of Scheduled Passenger Traffic* [online]. Available from: <http://www.icao.int/sustainability/pages/eap_fp_forecast_longterm.aspx>. [Accessed 5 April 2013].

ICAO - International Civil Aviation Organization (2013b). *Facts and Figures - World Aviation and the World Economy* [online]. Available from: <http://www.icao.int/sustainability/Pages/Facts-Figures_WorldEconomyData.aspx>. [Accessed 15 April 2013].

Kanafani, A. K. (1983). *Transportation demand analysis*. 1st. ed. United States: McGraw-Hill, Inc..

Krupka, M. (2010). *Last A310 leaves CSA fleet*. *Airplane pictures* [online]. Available from: <http://www.airplane-pictures.net/news.php?p=792>. [Accessed 4 February 2013].

LAWA - Los Angeles World Airports (2012). Statistics — Ten Year Summary — Passengers [online]. Available from: <http://www.lawa.org/welcome_LAX.aspx?id=800>. [Accessed 16 November 2012].

Lohatepanont, M. and Barnhart, C. (2004). *Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment*. Transportation Science. 38, pp.19-32.

Market Analyser (2013). *MIDT* [online]. Available from: <http://ma.aspirion.aero/midt>. [Accessed 15 April 2013].

Matsumoto, K. (2007). *International air network structures and air traffic density of world cities*. Transportation Research Part E [online]. 43, p.269–282. Available from: <http://www.sciencedirect.com/science/article/pii/S1366554506000822>. [Accessed 12 May 2012].

Michikian, K. and Hassiotis, V. (2002). *Cross-Atlantic Flights*. Athens, Greece: Hellenic vACC. Available from: <http://www.e-haf.org/PublicFTP/322/322/Manuals/IFR/CrossAtlanticFlights.pdf>. [Accessed 7 October 2012].

Ponikelska, L. (2011). *Letiste Praha Aims to Double Number of Long-Haul Flights, Official Says*. Bloomberg [online]. Available from: <http://www.bloomberg.com/news/2011-08-12/letiste-praha-aims-to-double-number-of-long-haul-flights-official-says.html>. [Accessed 4 February 2013].

Powell, R. A., Mwakalonge, J. L. (2012). *Productivity Performance of US Passenger Airlines since Deregulation*. In: The Transportation Research Board (TRB) 92nd Annual Meeting, 13-17 January 2013, Washington, D.C. pp. 1-22.

Sabre (2013). *Sabre Airline Solution* [online]. Available from: <http://www.sabreairlinesolutions.com/home/>. [Accessed 4 April 2013].

Skalicky, J. (2012). *Praha přijde na podzim o přímé letecké spojení s USA*. Czech Radio [online]. Available from: <http://www.rozhlas.cz/zpravy/domaciekonomika/_zprava/1073163>. [Accessed 4 December 2012].

Stadt Zürich (2012). Facts [online]. Available from: <http://www.stadt-zuerich.ch/content/portal/en/index/portraet_der_stadt_zuerich/zahlen_u_fakten.html>. [Accessed 16 November 2012].

TRB - Transportation Research Board (2007). *Airport Cooperative Research Program Synthesis 2-Airport Aviation Activity Forecasting*. (Project 11-03). Washington, D.C. Available from: <http://onlinepubs.trb.org/onlinepubs/acrp/acrp_syn_002.pdf>. Accessed: 28th March 2013.

PRG - Vaclav Havel Prague Airport, Ruzyne (2013a). *Destinations* [online]. Available from: <http://www.prg.aero/en/flight-info/destination-and-airlines/destinations/>. [Accessed 25 March 2013].

PRG - Vaclav Havel Prague Airport, Ruzyne (2013b). D*estination: Success - Let your story begin in Prague*. Prague: Prague Airport. Available from: <http://www.prg.aero/Files/letecky-obchod/brozura-letecky-obchod/>. [Accessed 30 March 2013].

PRG - Vaclav Havel Prague Airport, Ruzyne (2013c). *Business section* [online]. Available from: <http://www.prg.aero/en/business-section/>. [Accessed 30 March 2013].

PRG - Vaclav Havel Prague Airport, Ruzyne [online]. (2012). Available from: <http://www.prg.aero/en/>. [Accessed 4 October 2012].

PRG - Vaclav Havel Prague Airport, Ruzyne (2012b). *Unserved Routes – Map and MIDT data* [online]. Available from: <http://www.prg.aero/en/business-section/aviation-business/new-market-opportunities/unserved-routes-map-and-midt-data/>. [Accessed 13 April 2013].

Srinidhi, S. (2009). *Development of an Airline Traffic Forecasting Model on International Sectors*. In: 5th Annual IEEE Conference on Automation Science and Engineering, 22-25 August 2009, Bangalore, India. pp. 1-6.

U.S. Bureau of Labor Statistics (2012). Unemployment Rates for Metropolitan Areas [online]. Available from: <http://www.bls.gov/lau/lamtrk11.htm>. [Accessed 16 November 2012].

U.S. Census Bureau (2012). State [online]. Available from: <http://quickfacts.census.gov/qfd/states/>. [Accessed 16 November 2012].

US DOT - U.S. Department of Transportation (2013). Passengers All Carriers - All Airports [online]. Available from: <http://www.transtats.bts.gov/Data_Elements.aspx?Data=1>. [Accessed 23 February 2013].

Weber, M. and Williams, G. December 2001. *Drivers of long-haul air transport route development*. Journal of Transport Geography. 9, pp.243–254.]

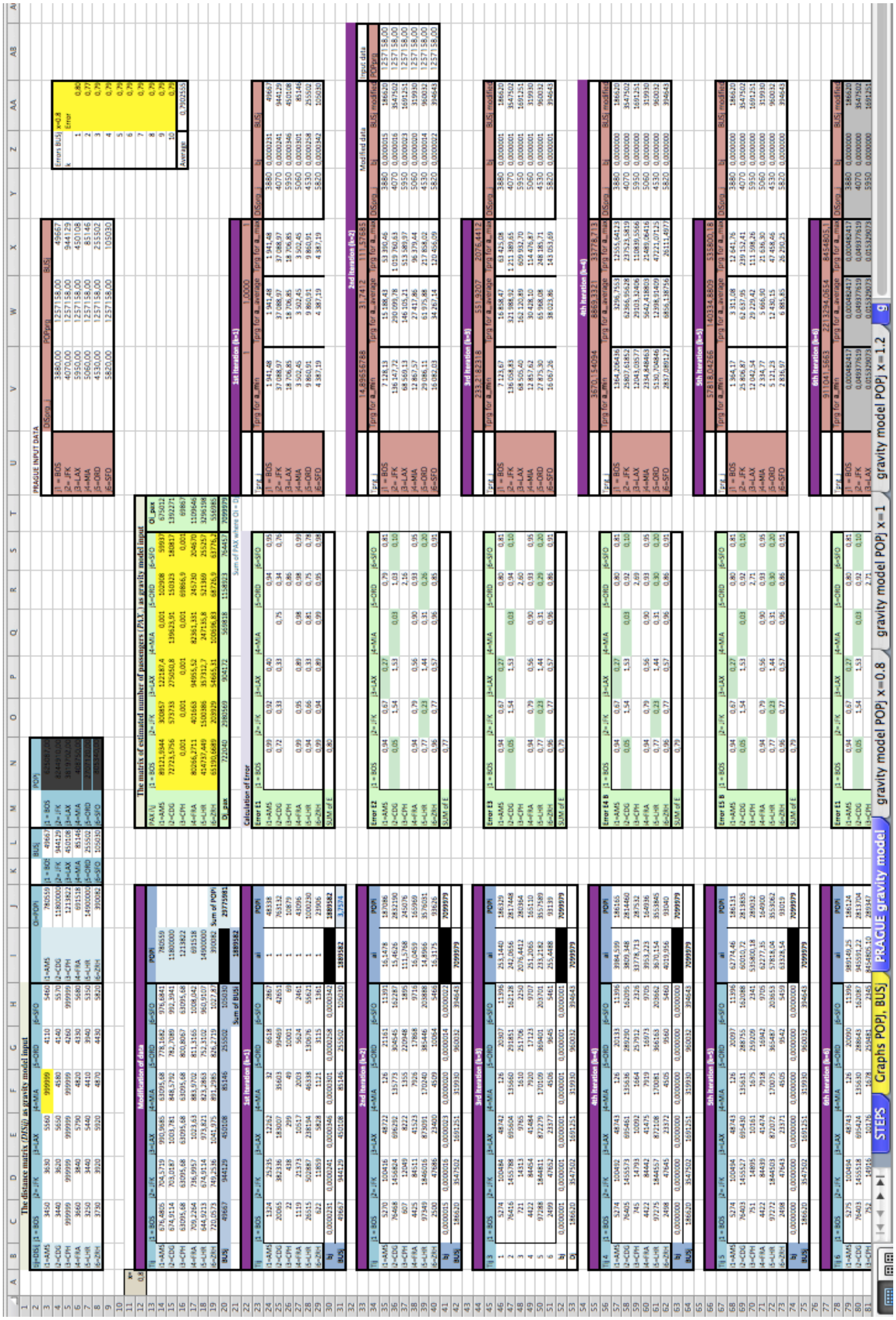
WebFlyer (2013). Mileage Calculator [online]. Available from: <http://webflyer.com/travel/mileage_calculator/>. [Accessed 12 February 2013].

Wensveen, J. G. (2007). *Air Transportation: A Management Perspective*. 6th. ed. Hampshire, England: Ashgate Publishing, Ltd..

Wien.at (2012). Statistik Wien [online]. Available from: <http://www.wien.gv.at/statistik/>. [Accessed 16 November 2012].

Zurich Airport (2012). About the airport [online]. Available from: <http://www.zurich-airport.com/desktopdefault.aspx/tabid-31/>. [Accessed 16 November 2012].

Appendix A


STEPS
Graphs POPj, BUSj
PRAGUE gravity model
gravity model POPj x=0.8
gravity model POPj x=1
gravity model POPj x=1.2

Vita

Kateřina Střelcová was born in Hranice, Czech Republic. She hold Master`s degree of Transportation Engineering in Air Traffic Control and Management from 2012 and a four year Bachelor`s Degree in Air Transportation from the Czech Technical University in Prague. She also studied Intelligent Transport Systems and Aircraft Design as exchange student at the Institute of technology at the Linkoping University in Sweden. She has more than 3 years of experience in various projects in satellite technology including working on a project Satellite Navigation and Practical Use in Transportation. Within this long-term project, she is also a co-author of the Annual report "Analysis of the technical and metrological requirements for equipment of GNSS receivers and their operations – 1. phase" (Prague CVUT FD, 2008). Her other projects include her bachelor thesis "Global navigation satellite systems applications certification methods for transport" and her master thesis "Application of APV SBAS for approach of aircraft at European airports"

Simultaneously, she has begun to study dual master degree program, ATLANTIS, in cooperation between Czech Technical University in Prague and University of Texas at El Paso. In fall 2012 she entered the College of Engineering at the University of Texas at El Paso to complete her second Master`s degree in Engineering in May 2013. During this period she worked on her master thesis that she also presented on spring Texas District ITE (Institute of Transportation Engineers) meeting in 2013.

Moreover, she gained working experience as Project Manager for State environment fund of the Czech Republic. In addition, She has around three and a half years of experience with managing Integrated Management Systems based on International Organization for Standardization ISO 9001, ISO 14001 and Occupational Health and Safety Advisory Services OHSAS 18001 in a logistics company as Manager of Quality, Safety, and Environment.

Permanent address: Svat. Cecha 788

Lipnik nad Becvou, Czech Republic, 751 31

This thesis was typed by Kateřina Střelcová.