

## ABSTRACT

Title of Document: DETERMINING THE RELATIONSHIPS  
AMONG AIRPORT OPERATIONAL  
PERFORMANCE AREAS AND OTHER  
AIRPORT CHARACTERISTICS

Kennis Yuen Man Chan, M.S., 2009

Directed By: Professor David J. Lovell  
Department of Civil and Environmental  
Engineering

In this thesis, a methodology is proposed to investigate pair-wise relationships between different types of airport operational performance variables. The methodology represents a fundamental contribution for comparing airport performance between different air traffic management systems. Considerable attention is paid to analyzing the most appropriate techniques in an effort to produce the most reliable results. Additionally, a method to display the results in a simple and clear way is also suggested to allow users to understand the results visually. The key variables obtained from the proposed methodology not only serve as building blocks for developing models to answer a variety of air traffic questions, which allow policy makers to make decisions on allocating resources wisely, but also can be used as an evaluation tool to assist the FAA in selecting candidate projects.

DETERMINING THE RELATIONSHIPS AMONG AIRPORT OPERATIONAL  
PERFORMANCE AREAS AND OTHER AIRPORT CHARACTERISTICS

By

Kennis Yuen Man Chan

Thesis submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Master of Science  
2009

Advisory Committee:  
Professor David J. Lovell, Chair  
Professor Michael O. Ball  
Professor Paul M. Schonfeld

© Copyright by  
Kennis Yuen Man Chan  
2009

## Acknowledgements

First of all, I would like to sincerely thank my advisor Professor Lovell for his guidance of my Masters thesis. Professor Lovell helped me to develop my critical thinking skills that will be beneficial to me in the future. I also like to thank for Professor Ball for his advice in both my project work and my thesis work. Their patience and support has been a great help for me to my success in graduate school. Their knowledge in air traffic research has brought me to a better understanding with air traffic research. Without their support, my goal of getting a masters degree would not be possible. I am also deeply grateful to Professor Schonfeld for his kindness and encouragement during my school years in UMD. Thanks to Prof. Lovell, Prof. Ball and Prof. Schonfeld for serving on my thesis committee.

Thank you to all my friends in the NEXTOR group for making my time at UMD enjoyable. The advice is helpful and the cakes are delicious.

Friends from ATAC Corporation also played critical roles in the completion of my masters degree. Thanks to ATAC for being so supportive for my decision of coming back to school and offering help when I needed it.

Finally, I must thank Dr. Tomonori Honda, Julia Law, YoonHee Shim and my mom for being such great listeners. Thanks for Tomonori especially for encouraging me to accept new challenges and helping me to adapt to the new environment in Maryland and my status as a graduate student. My time on the east coast would have been completely different without you.

## Table of Contents

Acknowledgements .....	ii
List of Tables .....	v
List of Figures .....	vi
Chapter 1: Introduction.....	1
<u>1.1 Motivation for determining the relationships among airport performance areas and airport characteristics</u> .....	2
<u>1.2 Organization</u> .....	7
Chapter 2: Airport operational performance variables.....	9
<u>2.1 Airport Delay</u> .....	9
<u>2.2 Airport Utilization</u> .....	11
<u>2.3 Cancellation Ratio</u> .....	12
<u>2.4 Airport Throughput</u> .....	13
<u>2.5 Airport Capacity</u> .....	14
<u>2.6 Airport Hub Structure and Multiple Airport System (MAS)</u> .....	15
<u>2.7 Number of runways</u> .....	16
Chapter 3: Data sources and preparation.....	16
<u>3.1 Aviation System Performance Metrics (ASPM) Individual flight data</u> .....	17
<u>3.2 Bureau of Transportation statistics (BTS) data</u> .....	19
<u>3.3 Enhanced Traffic Management System (ETMS) data</u> .....	19
<u>3.4 Other Airport characteristics data</u> .....	19
3.4.1 Runway information .....	20
3.4.2 Hub and Multiple Airport System information .....	21
<u>3.5 European Data</u> .....	21
3.5.1 European Flight Data .....	21
3.5.2 European Runway Information.....	23
<u>3.6 Data discrepancies</u> .....	24
Chapter 4: Methodologies.....	27
<u>4.1 Metrics for similarity to cluster airports</u> .....	29
4.1.1 Proposed metrics to cluster airports using hourly delay profiles.....	30
4.1.2 Proposed metric to cluster airports using other airport operational performance variables.....	33
<u>4.2 Hierarchical clustering procedures</u> .....	34
<u>4.3 Optimal number of clusters for nonhierarchical clustering procedures</u> .....	36
<u>4.4 Nonhierarchical clustering procedures</u> .....	38
<u>4.5 Statistical significance test</u> .....	41
4.5.1 Contingency Table.....	41
4.5.2 Fisher’s Exact Test.....	42
Chapter 5: Quantitative Analysis .....	44
<u>5.1 Design Structure Matrix (DSM)</u> .....	44
<u>5.2 Results and Data Interpretation</u> .....	46
<u>5.3 Results Validation</u> .....	54
Chapter 6: Conclusions.....	57
Chapter 7: Future Work.....	59

Appendices..... 60

## List of Tables

Table 1: Number of Runways[11].....	20
Table 2: Lists of Hub and MAS Information[1].....	21
Table 3: Field name and description in CODA and CFMU data.....	22
Table 4: European Runway Information[12].....	23
Table 5: Table illustrates the discrepancy in traffic counts between FAA databases.	26
Table 6: Hourly weights for delays.....	33
Table 7: Contingency Table - Sample.....	42
Table 8: ASPM carriers.....	61
Table 9: ASPM airports.....	62

## List of Figures

Figure 1: Delay statistics for U.S. and Europe.....	5
Figure 2: Cumulative Distribution Function for departure operations per runway per day between U.S. and Europe.....	5
Figure 3: Flowchart of the data process.....	7
Figure 4: Sample delay profile.....	11
Figure 5: Traffic count discrepancies in different FAA database.....	25
Figure 6: Total of departure operations using all 35 OEP airports combined in different databases/data fields.....	25
Figure 7: Illustration the sensitivity of using different similarity metrics.....	30
Figure 8: Comparisons of distance measures for single linkage and complete linkage [14] .....	35
Figure 9: Graph illustrates the chaining phenomenon in Single Linkage procedures	36
Figure 10: Cluster balance metric vs. using 3 months worth of data in 2006 .....	37
Figure 11: Cluster balance metric using one-year worth of data in 2006.....	38
Figure 12: K-means clustering algorithm Decision Diagram.....	39
Figure 13: Clustering results based on total airport utilization in 2008 .....	40
Figure 14: Airport utilization profiles colored by clusters .....	40
Figure 15: Sample P values table should the pair-wise relationships between airport operational performance variables (Red=P value less or equal to 0.01, indicate strong relationship). .....	43
Figure 16: DSM representation of the results using data from 2002-2008 .....	49
Figure 17: DSM representation of the results using 2002 data.....	50
Figure 18: DSM representation of the results using 2003 data.....	50
Figure 19: DSM representation of the results using 2004 data.....	51
Figure 20: DSM representation of the results using 2005 data.....	51
Figure 21: DSM representation of the results using 2006 data.....	52
Figure 22: DSM representation of the results using 2007 data.....	52
Figure 23: DSM representation of the results using 2008 data.....	53
Figure 24: DSM representation of results using yearly data from 2002-2008 .....	54
Figure 25: DSM representation of correlation analysis results using yearly data from 2002-2008 .....	56
Figure 26: Description of variables.....	60



## Chapter 1: Introduction

Although the overall air traffic demand is expected to grow slower than previously forecast for the near future, the FAA continues to forecast a long-term overall aviation growth in the United States despite challenges caused by the recent economic downturn over the world. Demand at some airports is increasing at an alarming rate. One example is the 20 percent flight operations increase from 2006 to 2007 at New York's John F. Kennedy International Airport (JFK). Other examples are the recent New York airspace redesign project and the order to temporarily limit the number of scheduled operations at JFK by the FAA. All the examples show that traffic congestion in the New York area is still a burning issue. The long-term overall aviation growth and the continued growth in demand at some already-congested domestic airports indicate that there is still a need to improve the existing air traffic management system in order to accept the on-going and future demands. In fact, the airport congestion problem is not only a domestic issue but also a worldwide issue. In Europe, most of the major hub airports are also experiencing heavy congestion concentrating at the peak time of the day. Both U.S. and Europe are actively looking for solutions to improve air traffic performance and on-going research efforts have been seen through the NextGen project in U.S. and SESAR in Europe.

Due to the similarities in the air traffic management (ATM) infrastructure in U.S and Europe, comparing the two ATM systems not only allows us to better understand the similarities and differences of the systems but most importantly it will help us to identify better ATM practices. The results will be greatly beneficial to both systems. Over the years, there are many high level statistics showing the differences

in the performance of the ATM systems in the U.S. and Europe individually. However, very limited analyses have been done to compare the performance of the two systems. In those analyses, only one or a few performance areas have been looked at. Except for the on-time performance measure, there is presently a lack of commonly agreed upon and comparable performance indicators worldwide [1]. Simple statistics and inconsistent performance measurements could lead to biased and incorrect conclusions. Therefore, to effectively and fairly compare the two ATM systems, the analysis has to be done in a coherent and consistent way. Before performing the comparison, we should have a fundamental understanding of how the air traffic systems work in both the U.S and Europe. Without sufficient knowledge of how the air traffic system behaves, the comparison will be invalid and results will not be accurate. This thesis intends to provide methodologies to help us to get the fundamental knowledge that we need to perform the U.S./Europe comparison in the future. The results of the methodologies also can help in answering a variety of air traffic questions with further analysis.

### *1.1 Motivation for determining the relationships among airport performance areas and airport characteristics*

In 2009, the FAA collaborated with EUROCONTROL to carry out a very detailed high-level comparison of airport operational performance between the U.S. and Europe Air Navigation systems [2]. Different airport operational performance areas were compared between the two systems by phase of flight. The paper pointed out many important operational performance similarities and differences between the two systems. This global comparison serves as a very good starting point for both

Europe and the U.S to understand the two systems in a more consistent way. In addition to high-level comparisons, the FAA also sees a need to perform a more advanced mathematical and statistical comparison between the ATM systems. As a result, the FAA funded the NEXTOR group to perform more detailed global comparisons, which could eventually help policy makers to make sound decisions to improve the current ATM systems.

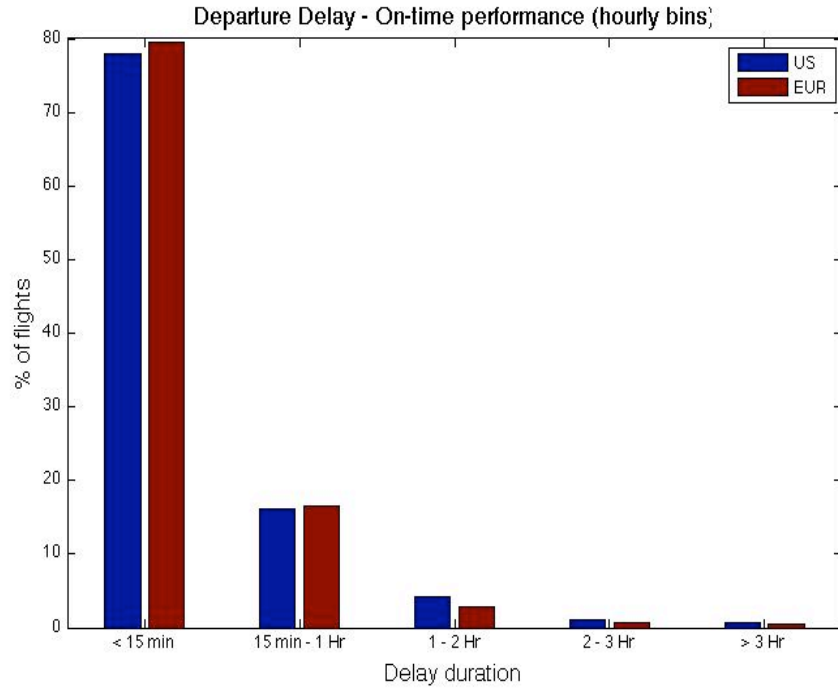
The NEXTOR group has access to various types of U.S airport performance data. Unfortunately, only a very small amount of test sample data from Europe has been received at this point. We cannot perform meaningful comparisons without sufficient data from both systems, because it is a data-driven process. While efforts are underway to improve the data gathering efforts from the European side, a good way to move forward is to investigate thoroughly the methodologies that can be employed (see Figure 3), and that is the subject of this thesis. At this point, the methodologies can only be demonstrated with a large set of U.S. data, but they have been designed with this more comprehensive comparison in mind for the future. The small amount of sample European data obtained so far from EUROCONTROL is used to illustrate the importance of the methodologies developed in this study.

To accurately compare two ATM systems, we should consider *what* to compare before considering *how* to compare. Airport operational performance is often measured by airport delay. However, is delay alone a good indicator to measure airport operational performance? A common metric for evaluating airport operational performance using flight delay information is airport “on-time performance.” Airport on-time performance is defined as the proportion of flights that arrive early or that are

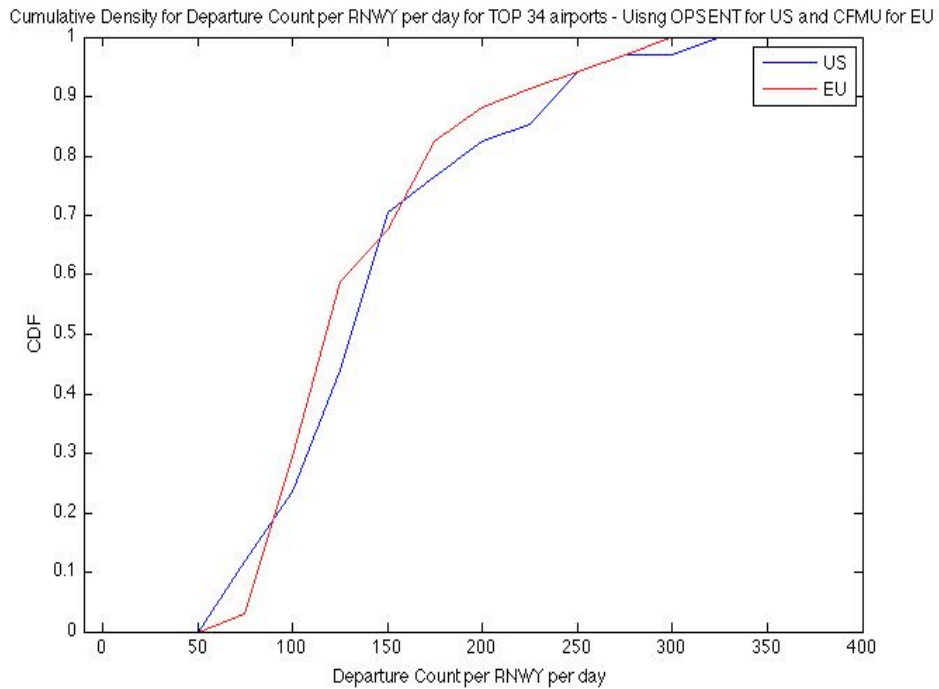
delayed by less than 15 minutes compared to their scheduled times. When we use on-time performance to compare airport performance between the ATM systems in Europe and the U.S.<sup>1</sup>, we observe the following phenomena: Europe has a slightly higher percentage of on-time flights (see Figure 1). However, when we analyze the number of actual daily operations per runway for each airport (see Figure 2), which represents airport utilization when capacity data are not available, European airports tend to handle fewer flights in general. Lower airport utilization could be one of the main reasons why European airports have better on-time performance. Therefore, when we evaluate airport operational performance, we must systematically analyze the delay and its interdependent (complementary) factors together, rather than just airport delays alone. This study provides methodologies to determine the “links” between airport performance variables such as airport delays, utilization, cancellation ratio, etc. Once the key variables are identified, more advanced evaluations can be performed. The results of this study have various important applications, especially in model development and model validation, and some of these applications will be explained in detail in the concluding section of this thesis.

---

<sup>1</sup> Currently, the European data is limited to five days.



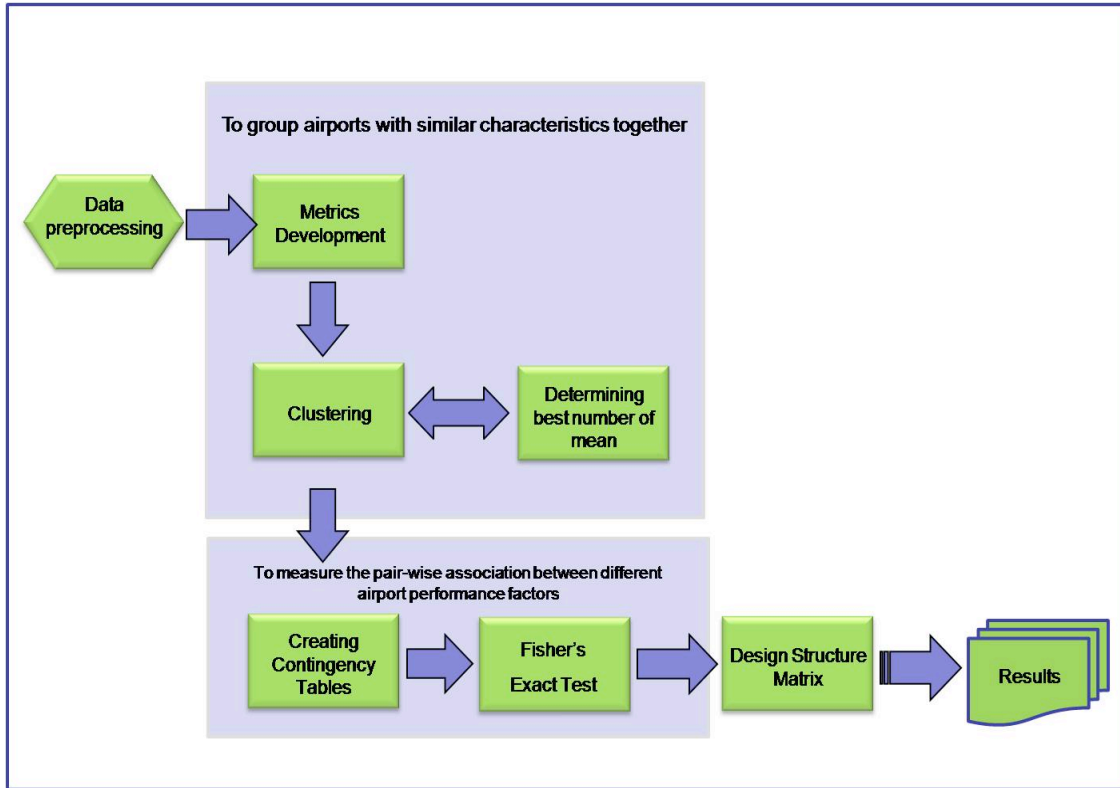
**Figure 1: Delay statistics for U.S. and Europe**



**Figure 2: Cumulative Distribution Function for departure operations per runway per day between U.S. and Europe**

Only major airports in the U.S. will be evaluated based on a group of operational performance areas identified by the FAA and SESAR such as throughput,

capacity, utilization, efficiency, etc. However, factors that contribute to the differences in airport operational performance will also be evaluated. All the areas and factors being evaluated are called airport operational performance variables in this study. The results of this study can be used to improve an airport operational performance variable by adjusting the other airport operational performance variable(s) with strong relationships. For example, we can use the results to find out the main factors causing flight delays at airports. By carefully and skillfully adjusting the main factors for flight delays, flight delay at the airports can be improved. The results of this study also provide significant information used in statistical modeling for airport delay prediction. Most importantly, we can use the results to select the main performance variables in the modeling process to compare the airport operational performance between the U.S and Europe once performance variables and similarity metrics for clustering analysis are defined.



**Figure 3: Flowchart of the data process**

## 1.2 Organization

This document contains seven chapters. The first chapter introduces the motivation of this study and why it is important to understand the pair-wise relationships between different airport operational performance variables. It also provides a flow diagram as a basic overview of the method to reveal the pair-wise relationships, which is the main theme of this study. The second chapter identifies some of the commonly used airport performance variables and their definitions. The third chapter lists all the data sources and how new fields are derived for this study. Chapter four provides the procedures and techniques of the proposed methodologies. This chapter addresses the advantages and disadvantages of different relative techniques and provides explanations of why a certain technique is chosen over another. We also provide remedies for the disadvantages of the chosen techniques in an effort to produce

reliable results. Chapter five shows some quantitative analysis of the outputs from the proposed method and then final results are presented. The conclusions of this study and possible applications of the proposed method are discussed in chapter six. The final chapter also proposes some ideas on how this study can be extended in the future.



## Chapter 2: Airport operational performance variables

This chapter contains information on the formal definitions and rationale behind our choices of different airport operational performance variables. These airport operational performance variables could be defined differently depending on the purpose of the study and the availability of the data. In this thesis, we are providing several ways as examples to look at each variable to see how each one relates to the others. The definitions of each variable can be changed easily if necessary. The list of variables below should be expanded to include some other variable such as aircraft mix to cover all the airport performance factors when we analyze the ATM system performance between Europe and the U.S. The proposed method is designed to investigate pair-wise relationships between variables, therefore, excluding or including extra variables will not affect the reliability of the results. However, if we are uncertain of the relationship of a variable to the rest of the airport performance variables, we should always include it into the analysis and let the proposed method determine the relationships for us.

We are focusing on performance variables related to individual airport comparisons in this thesis. However, we should extend this analysis with system level variables such as propagated delays in the future because some performance differences might be a result of the network structure of the system.

### 2.1 Airport Delay

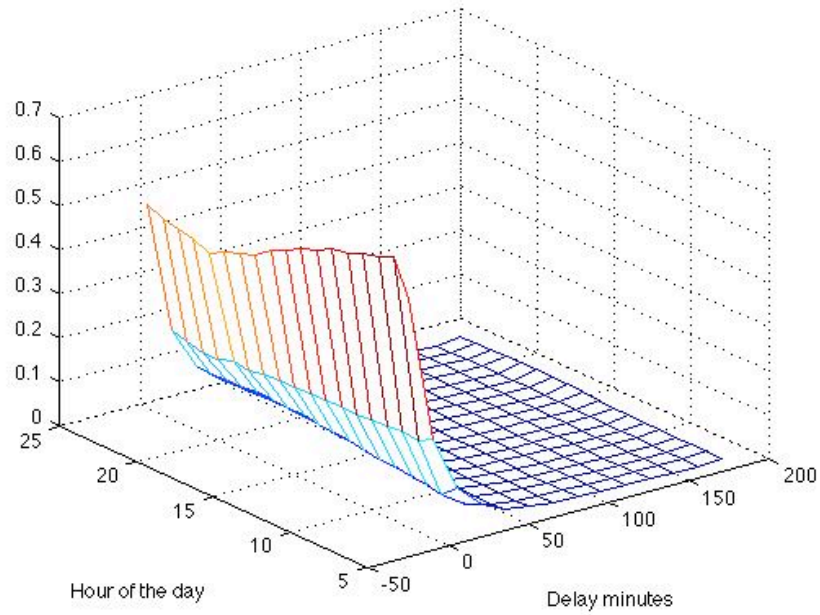
Airport delay is calculated as the difference between scheduled gate time and actual gate time. One of the most common ways to capture delay is using the average

delay. This can be a misleading statistic, however, because airports with the same number of delayed flights or the same average delay do not necessarily have the same performance. The time of day at which the delays occurred, for example, can have a significant impact to the system, especially since most of the 35 OEP airports act as hubs. Thus, it would be better to analyze the hourly delay distribution rather than average delay.

In this study, each hourly delay distribution is captured by a probability density function (see Figure 4) using 15 minute time bins. The details of how the 15 minute time bins are created can be found in Chapter 3. The matrix representation of the delay probability density function for a given day is:

$$Dp = \begin{pmatrix} p_6(1) & \cdots & p_6(n) \\ \vdots & \ddots & \vdots \\ p_{23}(n) & \cdots & p_{23}(n) \end{pmatrix} \quad (1)$$

where  $[p_i(1), \dots, p_i(n)]$  represents a vector of discrete probability density for hourly delay at the  $i^{\text{th}}$  hour of the day, and  $n$  is the bin number. This delay formulation will be applied to both arrival and departure delay using gate arrival and gate departure time respectively.



**Figure 4: Sample delay profile**

## 2.2 Airport Utilization

Airport utilization is calculated by the ratio of the number of scheduled flight operations to the airport capacity.

$$U_i = \frac{(ARR + DEP)_i}{(AAR + ADR)_i} \quad (2)$$

In this equation, the notation *ARR* represents the total number of scheduled arrival flights and *DEP* represents the total number of scheduled departure flights, using gate arrival and gate departure time. The airport-supplied departure rate (*ADR*) and the airport-supplied arrival rate (*AAR*) are used as the source to capture airport capacity provided by the FAA. Utilization should be calculated for each time interval ( $U_i$ ), because delays are very sensitive to changes in demand and capacity [3] when an

airport operates close to its maximum capacity. Capacity values are recorded per hour in the data; therefore, it would be appropriate to also calculate utilization in one-hour intervals. This matches with the discretization of delay into hourly probability densities. Utilization is one of the major airport performance variables because it tells us how busy the airport is at a given time. In this study, we will evaluate arrival utilization, departure utilization and total airport utilization. What we want to find out is if there are any different correlations with utilization and other airport performance variables when we analyze the arrival and departure data separately.

When we perform the European and US comparison study in the future, we should modify the definitions for utilization to include unscheduled operations such as general aviation. The U.S handled 4.5 times [4] more general aviation than Europe in 2007 (Europe had a 4% share of general aviation and the U.S had an 18% share of general aviation). This significant difference in general aviation could be one of the reasons for the performance differences between two systems.

### 2.3 Cancellation Ratio

Cancellation Ratio is calculated by the ratio between the number of cancelled flights and the number of gate operations. Cancellation ratio is also captured on an hourly basis because more cancellations in peak hours could have a different impact on the airport performance than cancellation in off-peak hours. We include cancellation ratio as one of the variables here because we believe that high flight cancellation ratios, especially at the peak hours, could have significant impacts to the air traffic system.

## 2.4 Airport Throughput

Airport throughput is the traffic volume at an airport over a period of time. It is one of the important determinants in evaluating airport performance because it directly ties to the changing of the delay distributions. One popular way to capture throughput at an airport is to look at the hourly throughput distribution throughout the day. If an airport operates at its maximum throughput for an extended period of time, delays will reach unacceptable levels. Rather than using day as the unit of time, hourly airport throughput is calculated by determining the total operations using wheels on and wheels off time.

The other useful way to evaluate airport performance using throughput is to look at the hourly throughput variability vector over time. Hourly throughput variability  $(TP_{COV})_i$  over a period of time, such as a year or a quarter, is computed by the Coefficient of Variation (COV) of throughput for each hour  $i$ . By looking at the hourly throughput COV, we can determine how traffic fluctuates for a particular hour over time for each hour. We have chosen hourly throughput variability because this hourly throughput should capture the consistency of usage of airports. The more consistent the airport usage is, the more easily the airport can be managed.

$$(TP_{COV})_i = \frac{\sigma_i}{\mu_i} \quad (3)$$

$$TP_{COV} = [(TP_{COV})_6, \dots, (TP_{COV})_{23}] \quad (4)$$

where  $\sigma_i$  is the standard deviation and  $\mu_i$  is the mean of the throughput.

Other options to look at throughput variability are:

- Daily Coefficient of Variation that captures variability of the throughput within a day.
- Frequency of throughput variation over the day captured by determining the number of local peaks in each day. Queuing theory suggests that this pattern will influence delays.

The above two options are not included in this study; however, it can be easily incorporated into the analysis in the future.

### 2.5 Airport Capacity

In this study, we are using the AAR and ADR values published by the FAA as input for capacity. AAR and ADR declare capacity based on airport infrastructure, weather conditions, runway configuration, and fleet mix [5]. The FAA updates the AAR and ADR values during the day when additional records are received. The AAR and ADR do not capture the maximum number of aircraft that can be handled by an airport. Instead, they capture the number of aircraft that the management thinks the airport can handle under specific conditions of that hour. Other than only capturing the hourly mean capacity over time, it is even more important to capture the variability of capacity throughout the day  $(C_{COV\_D})_d$  over time as well as the hourly capacity variability over a period of time. These variability measures are calculated in a manner similar to that described by Equations (3) and (4):

$$(C_{COV\_D})_d = \frac{\sigma_d}{\mu_d} \quad (5)$$

$\sigma_d$  is the standard deviation and  $\mu_d$  the mean for day  $d$ .

## 2.6 Airport Hub Structure and Multiple Airport System (MAS)

The hub type is considered as one of the performance variables because hubs for major carriers are more efficient than those that are not [1]. Following the FAA's hub definitions, hub type is classified by the percentage of national annual passenger boardings. A large hub has at least 1% of U.S. passenger boardings; medium hubs, more than 0.25% but less than 1%; small hubs, greater than 10,000 boarding passengers but less than 0.25%; and non-hub airports, at least 2500 boardings but no more than 10,000 boarding passengers[6]. We used the hub list from the FAA in this study because this list is widely used in other research projects. However, there may be a need to update the list with more recent data because the percentage of passengers boarding at airports changes over time. The grouping percentage used to classify the type of hub airports may also need to be changed to re-identify the hub structure better. One possible way to re-identify hub structure is to perform sensitivity analysis to hub definition in order to identify the most suitable percentage boarding passengers for hub classifications.

Sarkis [1] pointed out that since MAS airports, typically, have more passenger enplanements due to their locations in density populated areas, this will very likely increase the airport efficiency. He believes that airports in Multiple Airport Systems are more efficient than those in Single Airport Systems. Therefore, we include the MAS classification as one of the airport performance variables to indirectly capture airport efficiency. The hub and MAS airport list can be found in Chapter 3, Table 2.

### 2.7 Number of runways

Obviously, airport infrastructure may impact airport operational performance. However, airport infrastructure such as runway configurations cannot be easily quantified and, therefore, it will not be analyzed by the proposed methods in this study. One other way to differentiate infrastructure between airports is to quantify the number of runways at each airport. Recently, there have been some new runways opened at a few metropolitan airports and the new runway implementation year has been incorporated into this analysis.

## Chapter 3: Data sources and preparation

The data used in the study are mainly coming from three databases, the Aviation System Performance Metrics (ASPM) database, the Bureau of Transportation Statistics (BTS) database and the Operations Network (OPSNET) database. All three databases provide detailed flight information as well as basic statistics for the air transportation community to perform air transportation research and analysis; however, flight data in each of the databases are collected and processed differently to fit different needs. In this study, different fields in the three databases for the 35 Operation Evolution Plan (OEP) airports from 2002 to 2008 are extracted to capture different factors impacting airport operational performance. Some other quantifiable airport characteristics such as airport hub structure and number of runways are also included in this study, even though they are not one of the airport operational performances areas identified by the FAA or EUROCONTROL.



### 3.1 Aviation System Performance Metrics (ASPM) Individual flight data

The ASPM database includes records for the vast majority of commercial flights for the ASPM airports (see Table 9 in the Appendix), and for ASPM carriers (see Table 8 in the Appendix) regardless of airport [7]. Records in the database are compiled from the Enhanced Traffic Management System (ETMS), the Out, Off, On, In (OOOI) data from the airlines and the BTS Aviation System Quality and Performance (ASQP) system. The detailed flight delay information for individual flights in the database is the main reason for making the ASPM database the main data source for this study.

Data extracted from the ASPM individual flight records are:

- Arrival and Departure Airport
- Date (Year, Month, Day)
- Local Hours (6 to 23 hours only)
- Scheduled Gate Departure Time
- Actual Gate Out Time
- Scheduled Gate In Time
- Actual Gate In Time
- Scheduled Wheels Off Time
- Actual Wheels Off Time
- Scheduled Wheels On Time
- Actual Wheels On Time

Data Extracted from the ASPM hourly records are:

- Arrival and Departure Airport

- Date (Year, Month, Day)
- Local Hours (6 to 23 hours only)
- Airport supplied Departure Rate
- Airport Supplied Arrival Rate

Flight scheduled time and actual time are used to calculate flight delay. To obtain the delay distribution for airports, delay information per flight is assigned to different time bins based on the delay duration. Early arrival or departure flights are assigned to an early bin and delayed flights will be assigned to different 15 minute delay bins appropriately. Flights with delays above 179 minutes (3 hours) are assigned to the last delay bin. All the flights in the ASPM individual flight database are considered as scheduled flights in this study. The airport supplied departure rate (ADR) and the airport supplied arrival rate (AAR) are the sources for airport capacity at different airports. Since ADR and AAR information are only reported on an hourly basis, changes in airport capacity over the course of an hour will not be considered in this study.

Rather than using gate information, airport throughput is captured by using wheels on and off time for flights. Airport throughput is one of the important metrics for measuring airport performance. It captures the hourly usage at airports. Using wheels on and wheels off time reflects the runway occupancy time more accurately because flights with similar gate in/out time could arrive to or depart from the runway at different times.

### 3.2 Bureau of Transportation statistics (BTS) data

Among all the different types of data in the BTS system, only the Airline On-Time Performance Data is used in this study. The Airline On-Time Performance data contains scheduled and actual departure and arrival times reported by certified U.S. air carriers that account for at least one percent of domestic scheduled passenger revenues [8]. The main reason for using BTS data is to obtain flight cancellation data at the 35 OEP airports.

### 3.3 Enhanced Traffic Management System (ETMS) data

Enhanced Traffic Management system (ETMS) data is only used to compare the traffic counts with the sample European dataset in this study. ETMS is developed by the FAA, who provides both software and data to mainly allow the Air Traffic Control System Command Center (ATCSCC), the Air Route Traffic Control Centers (ARTCCs), and major Terminal Radar Approach Control (TRACON) facilities to manage the flow of air traffic within the National Airspace System (NAS)[9]. The ETMS data downloaded from the FAA ASPM website contains traffic counts broken down by either user groups or equipment types. It is considered a good set for traffic counts because “ETMSC contains every flight record constructed”[10].

### 3.4 Other Airport characteristics data

Data regarding the number of runways, new runway implementation date, airport hub structure and multiple airport system are obtained from either government websites directly or websites using government published data. The information is summarized and listed below:

### 3.4.1 Runway information

**Table 1: Number of Runways[11]**

Airport Name	FAA Airport code	Number Of Runway (As of 2009)	New Runway Opening Date
Atlanta Hartsfield International	ATL	5	27-May-06
Boston Logan International	BOS	6	23-Nov-06
Baltimore-Washington International	BWI	4	
Cleveland Hopkins International	CLE	4	
Charlotte Douglas International	CLT	3	
Cincinnati-Northern Kentucky International	CVG	4	19-Dec-05
Washington Reagan National	DCA	3	
Denver International	DEN	6	
Dallas-Ft Worth International	DFW	7	
Detroit Metropolitan Wayne County	DTW	6	
Newark International	EWR	3	
Ft Lauderdale-Hollywood International	FLL	3	
Honolulu International	HNL	4	
Washington Dulles International	IAD	4	21-Nov-08
George Bush Intercontinental	IAH	5	
John F Kennedy International	JFK	4	
Las Vegas McCarran International	LAS	4	
Los Angeles International	LAX	4	
La Guardia	LGA	2	
Orlando International	MCO	4	
Chicago Midway	MDW	5	
Memphis International	MEM	4	
Miami International	MIA	4	
Minneapolis-St Paul International	MSP	4	
Chicago O'Hare International	ORD	7	21-Nov-08
Portland International	PDX	3	
Philadelphia International	PHL	4	
Phoenix Sky Harbor International	PHX	3	
Pittsburgh International	PIT	4	
San Diego International	SAN	1	
Seattle-Tacoma International	SEA	3	21-Nov-08
San Francisco International	SFO	4	
Salt Lake City International	SLC	4	
Lambert-St Louis International	STL	4	
Tampa International	TPA	3	

### 3.4.2 Hub and Multiple Airport System information

**Table 2: Lists of Hub and MAS Information[1]**

<b>Airport Name</b>	<b>State</b>	<b>ICAO airport Code</b>	<b>Hub Type</b>	<b>MAS</b>
Atlanta Hartsfield International	GA	ATL	L	No
Boston Logan International	MA	BOS	L	Yes
Baltimore-Washington International	MD	BWI	L	Yes
Cleveland Hopkins International	OH	CLE	M	Yes
Charlotte Douglas International	NC	CLT	L	No
Cincinnati-Northern Kentucky International	KY	CVG	L	No
Ronald Reagan Washington National	DC	DCA	L	Yes
Denver International	CO	DEN	L	No
Dallas-Ft Worth International	TX	DFW	L	Yes
Detroit Metropolitan Wayne County	MI	DTW	L	Yes
Newark International	NJ	EWR	L	Yes
Ft Lauderdale-Hollywood International	FL	FLL	L	Yes
Honolulu International	HI	HNL	L	No
Washington Dulles International	VA	IAD	L	Yes
George Bush Intercontinental	TX	IAH	L	Yes
John F Kennedy International	NY	JFK	L	Yes
Las Vegas McCarran International	NV	LAS	L	No
Los Angeles International	CA	LAX	L	Yes
La Guardia	NY	LGA	L	Yes
Orlando International	FL	MCO	L	No
Chicago Midway	IL	MDW	L	Yes
Memphis International	TN	MEM	M	No
Miami International	FL	MIA	L	Yes
Minneapolis-St Paul International	MN	MSP	L	No
Chicago O'Hare International	IL	ORD	L	Yes
Portland International	OR	PDX	M	No
Philadelphia International	PA	PHL	L	No
Phoenix Sky Harbor International	AZ	PHX	L	No
Pittsburgh International	PA	PIT	L	No
San Diego International	CA	SAN	L	No
Seattle-Tacoma International	WA	SEA	L	No
San Francisco International	CA	SFO	L	Yes
Salt Lake City International	UT	SLC	L	No
Lambert-St Louis International	MO	STL	L	No
Tampa International	FL	TPA	L	No

## 3.5 European Data

### 3.5.1 European Flight Data

The central office and delay analysis (CODA) data and the Air Traffic Flow and Capacity Management (CFMU) data are provided by EUROCONTROL to

support the U.S./European study. Both datasets contain individual flight data covering 40 major European airports. The CODA data, which is used to compute delay information in this study, is supplied by the airlines. Airport demand is based on CFMU data. The current voluntary supply of CODA data represents approximately 70% among the 40 airports. The data sample that we obtained from EUROCONTROL is extracted in the CODA and CFMU systems to suit our needs and limit the volume of data that needs to be transferred; therefore, it only contains the fields in Table 3. Data shown in blue are based on CODA and data shown in black are based on CFMU.

**Table 3: Field name and description in CODA and CFMU data**

<b>File name</b>	<b>Description</b>
ADEP	Airport of Departure (ICAO)
IOBT	Planned off-block time as indicated in the last ATC-plan sent to Initial Flight Plan Processing System (IFPS) to CFMU
AOBT_3	Actual Off-Block time as calculated by CFMU
STD	Scheduled Time of Departure (Off-block) as communicated to the passengers
ACTUAL_OUT	Actual Out Actual off-block time
ACTUAL_OFF	Actual Off Actual Take-Off Time
ADES	Airport of Destination (ICAO)
ARVT_1	Planned landing time at destination as communicated to IFPS in the ATC-plan
ARVT_3	Actual landing time as calculated by CFMU
STA	Scheduled Time of Arrival (In-Block) as communicated to the passengers
ACTUAL_ON	Actual landing time
ACTUAL_IN	Actual In-Block time

Departure delay is computed by the difference between actual off-block time (ACTUAL\_OFF) and scheduled time of departure (STD). Arrival delay is computed by the difference between actual in-block time (ACTUAL\_IN) and scheduled time of arrival (STA). For total number of operations at airports, we sum up all the CMFU data per airport.

The sample European data is used to illustrate the difference in delays and traffic counts between the two systems in the introduction section. To develop the proposed methodology, only U.S data is used due to the availability of the data.

### 3.5.2 European Runway Information

Similarly to how we obtained runway information for airports in the U.S, runway information for the 40 European airports are obtained using an online database. The number of runways for each of the European airports in our study is listed in Table 4:

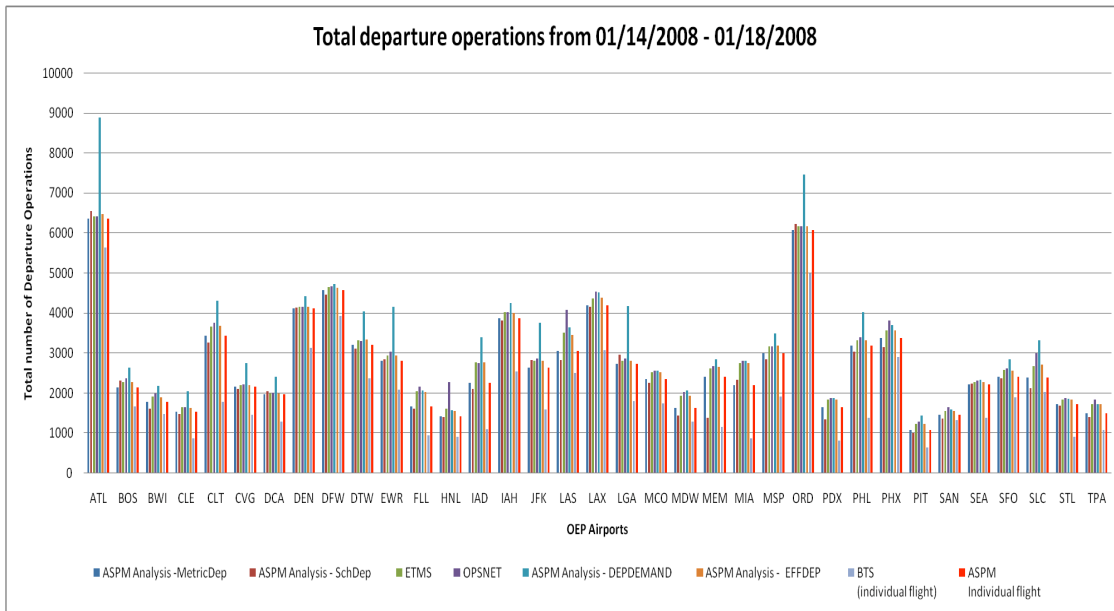
**Table 4: European Runway Information[12]**

ICAO Airport ID	Name	Country	Num of runway
EBBR	BRUSSELS NATL	Belgium	3
EDDF	FRANKFURT MAIN	Germany	3
EDDH	HAMBURG	Germany	2
EDDK	KOLN BONN	Germany	3
EDDL	DUSSELDORF	Germany	2
EDDM	MUNCHEN	Germany	2
EDDS	STUTT GART	Germany	1
EDDT	TEGEL	Germany	2
EFHK	HELSINKI VANTAA	Finland	3
EGCC	MANCHESTER	United Kingdom	2
EGGW	LUTON	United Kingdom	1
EGKK	GATWICK	United Kingdom	2
EGLL	HEATHROW	United Kingdom	2
EGPH	EDINBURGH	United Kingdom	2
EGSS	STANSTED	United Kingdom	1
EHAM	SCHIPHOL	Netherlands	6
EIDW	DUBLIN	Ireland	3
EKCH	KASTRUP	Denmark	3
ENGM	GARDERMOEN	Norway	2
EPWA	OKECIE	Poland	2
ESSA	ARLANDA	Sweden	3
LEBL	BARCELONA	Spain	3
LEMD	BARAJAS	Spain	4
LEMG	MALAGA	Spain	1
LEPA	PALMA DE MALLORCA	Spain	2
LFLI	SAINT EXUPERY	France	2
LFMN	COTE D AZUR	France	2
LFPG	CHARLES DE GAULLE	France	4
LFPO	ORLY	France	3
LGAV	ELEFTHERIOS VENIZELOS INTL	Greece	2
LIMC	MALPENSA	Italy	2
LIML	LINATE	Italy	2
LIRF	FIUMICINO	Italy	3
LKPR	RUZYNE	Czech Republic	2
LOWW	SCHWECHAT	Austria	2
LPPT	LISBOA	Portugal	2
LSGG	GENEVA COINTRIN	Switzerland	2
LSZH	ZURICH	Switzerland	3
LTAI	ANTALYA	Turkey	3
LTBA	ATATURK	Turkey	3

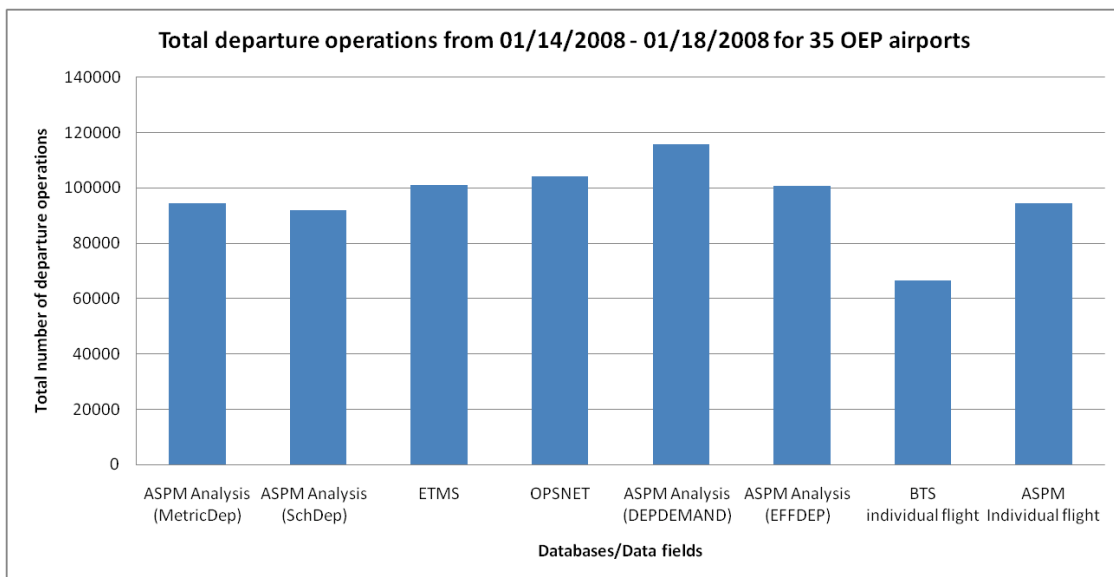
### 3.6 Data discrepancies

The FAA has developed various databases and metrics to provide the aviation community to access to historical traffic counts, forecasts of aviation activity, and delay statistics. Data in the FAA databases are collected and computed differently. Data in some databases, such as the Enhanced Traffic Management System (ETMS), are collected electronically and some are reported by the airports, such as the Operations Network (OPSNET). Due to the differences in the amount of data collected and the methods used to process the data in different databases, airport traffic statistics would be different (see Figure 5, 6 and Table 5). Hence, we need to carefully choose the type of data to use when evaluating airport performance. In this study, we use the individual flight ASPM database as the primary database to compute all the airport operational performance variables except cancellation ratio, even though there may be some other databases that can capture a certain performance variables better. The reasoning is that the ASPM individual flight database provides the most complete source among all other accessible data sources for our study. It contains “raw” data that allows us to derive different metrics for different airport performance variables. Most importantly, by using one central database, we can ensure that the errors are consistent among all the variables when finding the associations between them. We obtain cancellation ratios from the BTS database, as this is the only publicly accessible database that contains flight cancellation information.





**Figure 5: Traffic count discrepancies in different FAA database**



**Figure 6: Total of departure operations using all 35 OEP airports combined in different databases/data fields**

**Table 5: Table illustrates the discrepancy in traffic counts between FAA databases**

Total departure operations from 01/14/2008 - 01/18/2008								
Database	ASPM Analysis	ASPM Analysis	ETMS	OPSNET	ASPM Analysis	ASPM Analysis	BTS (individual flight)	ASPM Individual flight
Field Name	MetricDep	SchDep	OPSDEP	Airport Operations (Dep = Airport operations/2)	DEPDEMAND	EFFDEP		
ATL	6355	6554	6416	6414.5	8889	6468	5639	6355
BOS	2136	2318	2280	2376	2643	2270	1667	2136
BWI	1785	1615	1911	1989	2189	1904	1483	1785
CLE	1540	1476	1644	1642	2047	1641	876	1540
CLT	3432	3256	3670	3753	4300	3676	1778	3432
CVG	2154	2111	2204	2213	2742	2207	1463	2154
DCA	1972	2053	1990	2008	2401	1990	1298	1972
DEN	4119	4128	4157	4165	4412	4157	3138	4119
DFW	4571	4464	4640	4663.5	4717	4636	3922	4571
DTW	3201	3106	3312	3301.5	4046	3343	2373	3201
EWR	2815	2849	2946	3038	4160	2941	2078	2815
FLL	1664	1617	2054	2166	2061	2023	957	1664
HNL	1419	1403	1611	2277.5	1575	1565	916	1419
IAD	2259	2100	2774	2759.5	3393	2763	1101	2259
IAH	3869	3823	4016	4017	4252	4012	2537	3869
JFK	2639	2822	2805	2867.5	3765	2805	1599	2639
LAS	3062	2819	3510	4070.5	3649	3461	2500	3062
LAX	4185	4147	4371	4528	4518	4388	3081	4185
LGA	2738	2951	2808	2863.5	4167	2799	1812	2738
MCO	2344	2254	2522	2554	2563	2518	1754	2344
MDW	1635	1445	1934	2034.5	2064	1929	1287	1635
MEM	2402	1393	2625	2684.5	2851	2662	1149	2402
MIA	2196	2329	2750	2810.5	2806	2742	880	2196
MSP	2994	2850	3176	3168	3484	3186	1918	2994
ORD	6081	6228	6166	6166	7448	6176	4982	6081
PDX	1653	1347	1850	1884.5	1875	1842	808	1653
PHL	3192	3044	3325	3392.5	4028	3320	1385	3192
PHX	3379	3154	3564	3823.5	3699	3571	2908	3379
PIT	1080	999	1234	1286.5	1440	1230	640	1080
SAN	1470	1373	1553	1642	1585	1554	1320	1470
SEA	2222	2247	2280	2311.5	2336	2278	1384	2222
SFO	2403	2363	2573	2616.5	2843	2565	1895	2403
SLC	2398	2118	2684	2990	3316	2710	2038	2398
STL	1731	1688	1840	1870	1868	1838	903	1731
TPA	1496	1413	1726	1838.5	1735	1722	1079	1496

## Chapter 4: Methodologies

To accurately determine the relationships among different airport operational performance variables, a significant amount of data must be used to ensure that results are not biased and are legitimate for all airports in the system. Because each of the airports used in the study has some similarities, grouping airports based on their characteristics (which are the airport operational performance variables defined in Chapter 2), allows us to not only condense the data into manageable size but also reveal more concise and understandable descriptions of the data.

Furthermore, when we define the airport operational performance variables, there may be a need to look not only at a vector of an airport performance variable, but also at an array of probability distributions for an airport operational performance variable such as the delay profile (which is stored in a matrix). Matrix norms are numerous and varied, and none is used as popularly as correlation or covariance are used to describe vector differences. Instead, we will apply clustering analysis followed by statistical significance tests to identify associations between two airport operational performance variables.

In this study, we applied clustering analysis to group airports with similar properties. There are many different clustering techniques and each of them has its advantages and disadvantages. We have explored different techniques carefully and recognize the flaws of each of them. We provide remedies to tackle the shortcomings of the final chosen techniques to ensure that the proposed methodology produces the most reliable results among the methods we have examined.

Data clustering algorithms can be hierarchical or nonhierarchical. Since there are pros and cons of using either variety, we are using a mixed technique by combining both of them in this study to gain the benefits of each. For the hierarchical cluster procedures, a complete linkage agglomerative method is used. We decided to choose a complete linkage method because it produces better clustering results in this study than the other widely used hierarchical clustering method, called the single linkage method. As for the nonhierarchical procedures, a popular clustering technique called k-means clustering is applied. Other common clustering techniques are heuristic techniques such as tree-structure recursive partitioning and adaptive neural networks. However, these heuristic techniques are supervised learning techniques that require a priori knowledge of classification for the samples. Thus, these heuristic techniques are not applicable for this particular study.

One of the weaknesses of the k-means clustering as well as some other popular clustering techniques is that a metric to measure similarities and dissimilarities between objects (usually the Euclidean norm) must be determined ahead of time and the resulting solutions are highly sensitive to the defined metric. Conversely, some heuristic approaches will adjust the metric automatically. However, if the metric for similarity is defined properly, the results from the k-means algorithm can be explained more easily, compared to the heuristic approach. The detailed explanation of each of the similarity metric definitions used in k-means clustering can be found later in this chapter.

The other well-known disadvantage of using the k-means algorithm for clustering is that results could be very sensitive to the maximum number of clusters

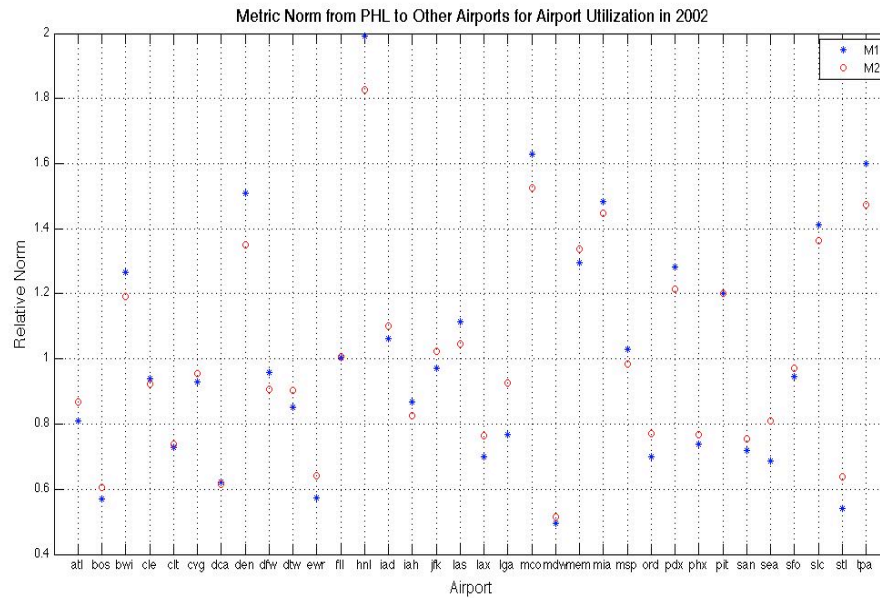
specified. One remedy for this problem is to employ a balancing method proposed by Jung [13] to locate the best number of groups during the airport clustering process.

Just like all other nonhierarchical clustering procedures, finding an appropriate way to initialize cluster seeds is one of the most complicated steps. Experiments show that the final cluster solution is very sensitive to the way we initialize the cluster seeds. Originally, the initial cluster seeds were selected randomly and different final cluster results were generated when re-runs showed that random starting seeds did not guarantee a consistent final solution. Therefore, rather than assigning the initial cluster seeds randomly, hierarchical cluster procedures with specific numbers of clusters are used to initialize the cluster seeds for the k-means algorithm.

#### 4.1 Metrics for similarity to cluster airports

As previously mentioned, the cluster solution is highly dependent upon the metrics used as the basis for similarity measures. To illustrate how sensitive the cluster solution changes by using different similarity measures, we can look at the changes in relative norm from one of the 35 OEP airports to Philadelphia International Airport (PHL) using two different metrics (see Figure 7). If the relative norm from an airport to PHL changes noticeably by using different metrics, the clustering results for that airport could be very different. There is no impact to Pittsburgh International Airport (PIT) when changing the metric to cluster airports based on airport utilization in 2002. On the contrary, the norm from LaGuardia Airport (LGA) to PHL changes notably when using different metrics. Therefore, selecting an appropriate metric to measure similarity becomes critical in clustering

analysis. In addition to some commonly used mathematical similarity metrics in clustering analysis (such as the Euclidean norm and the Manhattan norm), this study also provides other clustering metrics such as “weighted metric” and “average metric” which would be more appropriate for measuring similarities of certain airport operational performance areas or characteristics.



**Figure 7: Illustration the sensitivity of using different similarity metrics**

#### 4.1.1 Proposed metrics to cluster airports using hourly delay profiles

Considering that different similarity measures may lead to different cluster solutions and resulting in different conclusion, several options of similarity metric are provided to measure the norm between delay profiles (both arrival and departure delay) of airports in this study.

The four metrics are:

- Euclidean Norm
- Absolute Average delay

- Weighted Euclidean Norm
- Weighted Absolute Average delay

The metric for clustering to characterize airports based on the airport operational performance variables should depend on the data and what the air traffic community thinks the main feature(s) of the similarity metric should be. Unfortunately, there is no common agreement on what is the best metric to characterize airports. Therefore, other than using Euclidean Norm, which is one of the most traditional metrics for similarity measurement, we propose to use absolute difference between hourly average delay minutes. The Euclidean metric (Eqn. 5) and the absolute average delay metric (Eqn. 6) used in this study are as follows:

$$D_E(X, Y) = \sum_{i=6}^{23} \sqrt{\sum_{j=1}^K (X_{i,j} - Y_{i,j})^2} \quad (6)$$

where  $D_E(X, Y)$  represents the sum of the Euclidean norm (per hour) between two airport delay profiles, and  $X_{i,j}$  represents the probability associated with the  $j^{\text{th}}$  delay bin in the  $i^{\text{th}}$  hour at airport  $X$ .

$$D_A(X, Y) = \sum_{i=6}^{23} \left( \left| \sum_{j=1}^K T_j (X_{i,j} - Y_{i,j}) \right| \right) \quad (7)$$

where  $D_A(X, Y)$  represents the sum of absolute differences in average delay between two airports per hour and  $T_j$  represents the minimum delay associated with the  $j^{\text{th}}$  delay bin. Note that  $\sum_{j=1}^K T_j X_{i,j}$  represents average delay for the  $i^{\text{th}}$  hour.

Equations 6 and 7 assume that delay has the same impact to the system regardless of the time at which it occurs. However, this may not be a good

assumption because the queue formed due to delay during peak hours usually takes longer to dissipate. Therefore, a more realistic way to distinguish airports based on their delay profiles in the clustering process would be to assign different weights to the hours based on the time of the day when delay occurs. Thus, two additional delay metrics to measure the similarity of airport delay profiles are developed (Equations 8 and 9).

$$D_{WE}(X, Y) = \sum_{i=6}^{23} w_i \sqrt{\sum_{j=1}^K (X_{i,j} - Y_{i,j})^2} \quad (8)$$

Equation 7 is the same as Equation 5 except weights ( $w_i$ ) are assigned to different hours.

$$D_{WA}(X, Y) = \sum_{i=6}^{23} w_i \left( \left| \sum_{j=1}^K T_j (X_{i,j} - Y_{i,j}) \right| \right) \quad (9)$$

Similarly, Equation 8 is the same as Equation 6 except weights ( $w_i$ ) are assigned to different hours.

The weights for each hour are calculated by the percentage of total operations between 6:00a.m to midnight local time at the 35 OEP airports from 2002 to 2008 and they are listed in Table 6.



**Table 6: Hourly weights for delays**

Hour (local time)	Weight ( $w_i$ )
6:00-7:00	0.0331
7:00-8:00	0.0535
8:00-9:00	0.0642
9:00-10:00	0.0648
10:00-11:00	0.0643
11:00-12:00	0.0621
12:00-13:00	0.062
13:00-14:00	0.0639
14:00-15:00	0.0633
15:00-16:00	0.0638
16:00-17:00	0.0646
17:00-18:00	0.0645
18:00-19:00	0.0647
19:00-20:00	0.0636
20:00-21:00	0.0566
21:00-22:00	0.0456
22:00-23:00	0.0277
23:00-24:00	0.0178

4.1.2 Proposed metric to cluster airports using other airport operational performance variables.

Two popular metrics, the Manhattan norm (Eqn. 9) and the Euclidean norm (Eqn. 10), are used to measure the similarity for the rest of the airport operational performance variables suggested in Chapter 2.

$$D_e(\vec{x}, \vec{y}) = \sum_{i=6}^{23} |x_i - y_i| \quad (10)$$

$$D_r(\vec{x}, \vec{y}) = \sqrt{\sum_{i=6}^{23} (x_i - y_i)^2} \quad (11)$$

where  $\vec{x}$  and  $\vec{y}$  are vectors of airport operational performance for airport  $X$  and airport  $Y$  respectively and  $i$  represents the hour of the day.

The fundamental difference between the two metrics is that the Euclidean norm penalizes large differences between the two airport vectors more severely than the Manhattan norm. Therefore, we are providing two typical ways to measure similarity for all of the airport operational performance variables listed in Chapter 2 except delays.

#### 4.2 Hierarchical clustering procedures

The main function of hierarchical clustering in our study is to initialize the cluster seeds for the nonhierarchical clustering procedures. We do not use hierarchical clustering alone to do the clustering analysis here because hierarchical clustering methods do not guarantee optimal solutions, but the k-means algorithm guarantees local optima when it converges. Another advantage of using a hierarchical clustering method to choose the starting clusters is that the results from hierarchical clustering produces are deterministic and it is a clustered output coming out from a clustering method. The hierarchical clustering method used here is called the complete linkage method. Before using the complete linkage method, we have also evaluated another similar method called the single linkage method. The complete linkage method is similar to the single linkage method except that the cluster criterion is based on maximum distance rather than minimum distance (see Figure 8) [14].

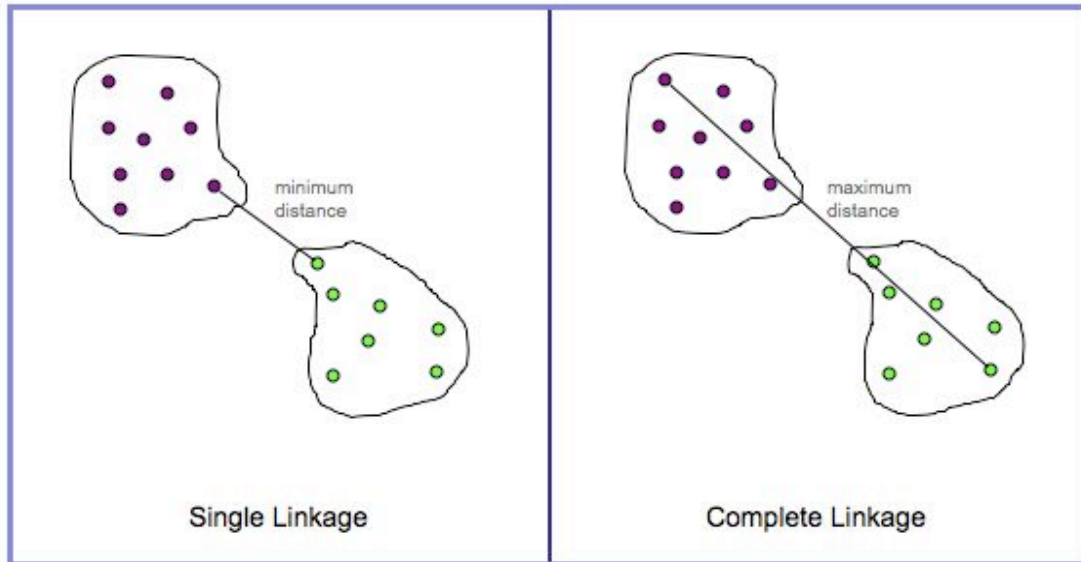
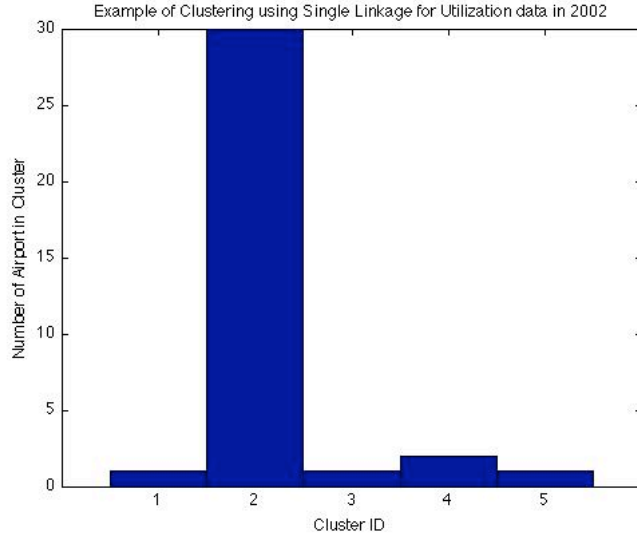


Figure 8: Comparisons of distance measures for single linkage and complete linkage [14]

One common phenomenon of the single linkage method is that clusters are forced to combine due only to two airports in different clusters being close to each other. This phenomenon is called “chaining phenomenon.” Due to the chaining phenomenon in the single linkage clustering method, the algorithm has a tendency to assign most of the clustering objects to the same cluster while leaving the distant objects to form new individual clusters by themselves (see example in Figure 9). Figure 9 shows that when we cluster the 35 OEP airports based on airport utilization in 2002 using the single linkage method, the majority of the airports are assigned to cluster 2 and distant airports formed new clusters by themselves. To avoid this chaining phenomenon in initial cluster seeds, the complete linkage method is used instead.



**Figure 9: Graph illustrates the chaining phenomenon in Single Linkage procedures**

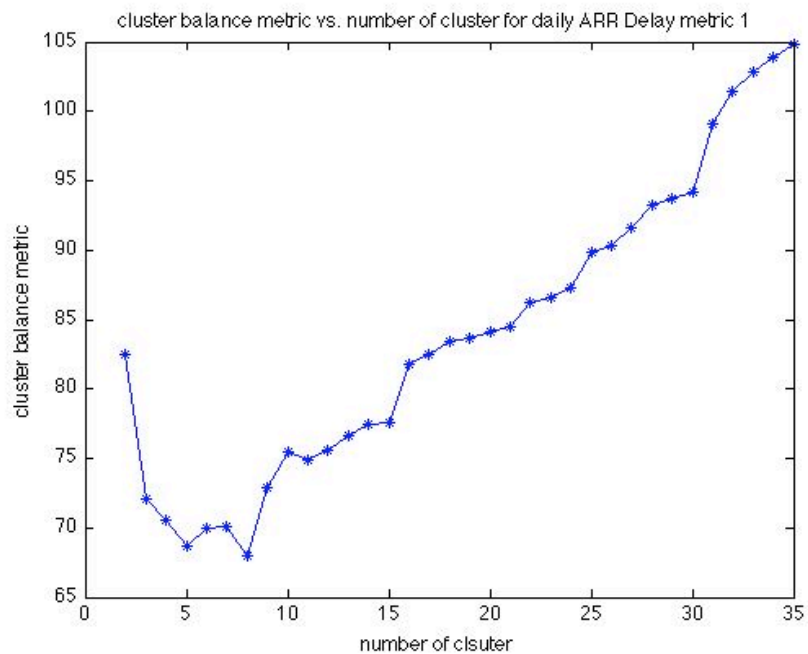
4.3 Optimal number of clusters for nonhierarchical clustering procedures

One of the drawbacks of using k-means clustering, as mentioned before, is that a maximum number of clusters would need to be chosen ahead of time and an inappropriate choice of the number of clusters could possibly yield suboptimal clustering results. To resolve this issue, we applied a method (see Eqn. 12) that is similar to the balancing technique suggested by Jung [13] to find the optimal number of clusters (OC) for k-means clustering. Jung’s balancing method uses clustering gain as a measure for clustering optimality. Although Jung applied his technique to a hierarchical clustering algorithm, his technique can be utilized to find optimal clusters for non-hierarchical clustering algorithms as well.

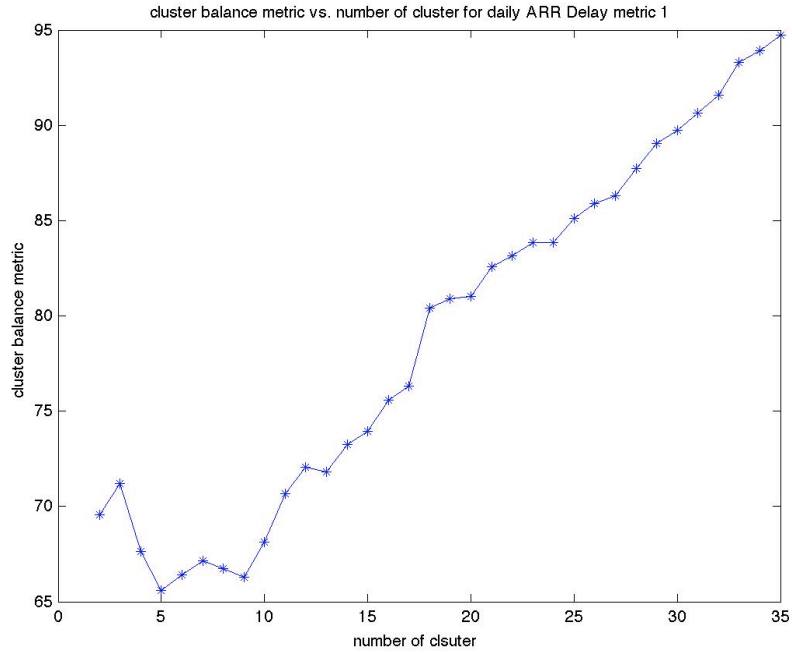
$$OC = Min \left[ \sum_j (\sum_i d(x_{i,j}, x_{j0})) + \sum_j d(x_{j0}, x_0) \right] \tag{12}$$

where  $x_{i,j}$  is the  $i^{th}$  element (airport) in cluster  $j$ ,  $x_{j0}$  is the center of cluster  $j$ , and  $x_0$  is the “overall” center of all samples.

The balancing method optimizes the number of clusters by compromising between the intra-cluster distance ( $d(x_{i,j}, x_{j_0})$ ) and the inter-cluster distance ( $d(x_{j_0}, x_0)$ ). To produce reliable results, the data need to be representative of the airports. Thus, enough data must be used to avoid clustering on the outliers of the data instead of the “real” trend of the data. As you can see in Figures 10 and 11, the optimal cluster number is dependent on the size of the data we use. Therefore, if we want to analyze quarterly trends, we should obtain the optimal cluster number for each year and then apply it to the quarters of that year to minimize the errors from the outlines. To look at the “overall” results without dividing the data into any subgroups, the optimal number of clusters is obtained using all seven years’ worth of data from 2002 to 2008. In this thesis, we will be looking at the yearly and overall results only.



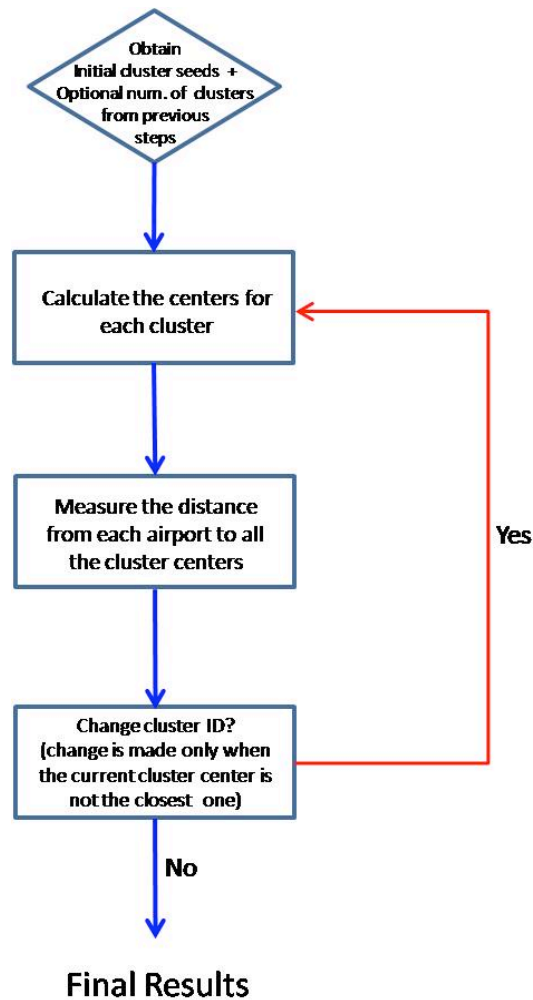
**Figure 10: Cluster balance metric vs. using 3 months worth of data in 2006 (Optimal number of cluster = 8)**



**Figure 11: Cluster balance metric using one-year worth of data in 2006 (Optimal number of cluster = 5)**

4.4 Nonhierarchical clustering procedures

Once the optimal number of clusters and good initial cluster seeds are determined, k-means clustering is used to “fine-tune” the results even further. Typically, the Euclidean norm is used as a metric for k-means clustering. However, depending on the meaning of “similarity” to the users, other metrics suggested in Section 4.1 could provide better descriptions of similarity. No matter what similarity metric we use, the k-mean algorithm procedure is the same. The procedures are listed in Figure 12.



**Figure 12: K-means clustering algorithm Decision Diagram**

After the K-means clustering analysis is performed, the 35 OEP airports will be clustered into different groups based on different performance variables. For example, when we cluster the 35 OEP airports based on total airport utilization, all the New York airports (LGA, JFK and EWR), PHL, ATL and ORD are clustered in the same group, as shown in Figure 13. It is because they are highly utilized throughout the day as seen in Figure 14. On the other hand, low utilization airports (airports colored in red) are clustered into the same group.

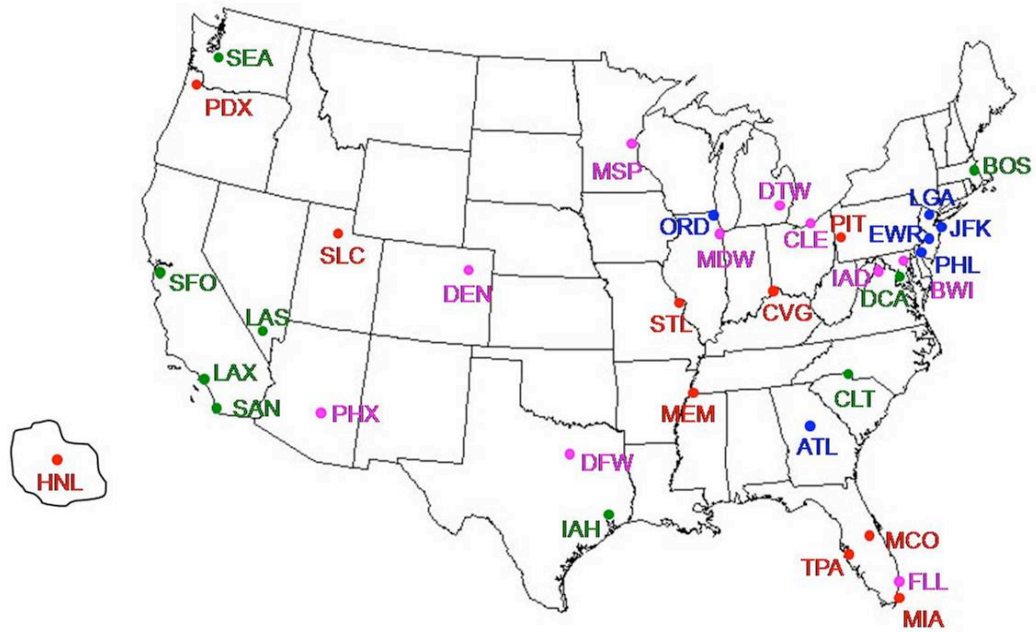


Figure 13: Clustering results based on total airport utilization in 2008

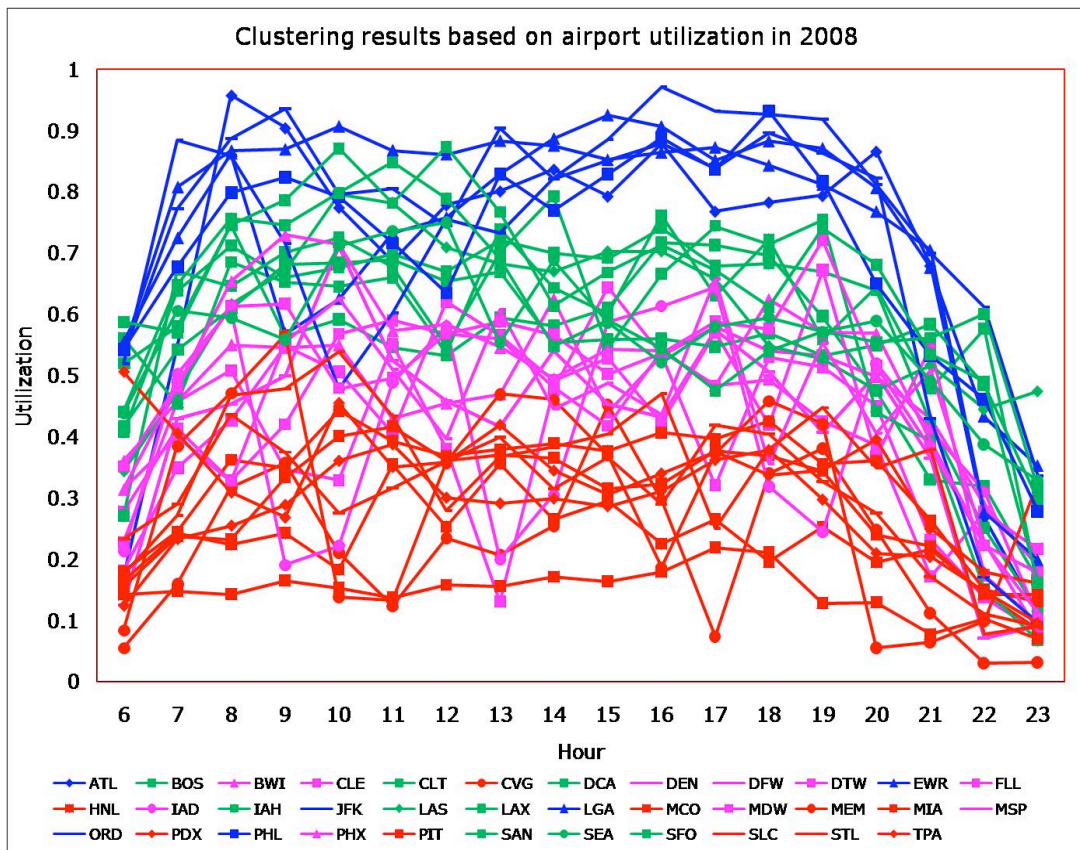


Figure 14: Airport utilization profiles colored by clusters



#### 4.5 Statistical significance test

Grouping airports is the first step of the process of finding pair-wise relationships between airport operational performance variables. After classifying the 35 OEP airports into different groups, we need to transform the information in meaningful ways. The main purpose of clustering is to analyze and determine which performance variables are inter-related. With the given cluster results, one way to determine a degree of relationship is by calculating associations between two variables.

##### 4.5.1 Contingency Table

Before we can determine the association between two clusters, information must be preprocessed into proper formats. Clustering of the performance variables converts data into categorical variables where each cluster represents a category. The Contingency Table is a common representation used to investigate relationships between two or more categorical variables. By putting the data into this format, relationships between two categorical variables can be visually understood.

Table 7 is an example of a contingency table for airport operational variable 1 and airport operational variable 2 for six airports. In this simple example, the optimal cluster number for both variables is two. Therefore, it is a 2x2 contingency table (the red box in Table 7). The total value in each row and column represents the number of airports for each cluster. The numbers in the cells of a contingency table represent the number of airports in common between the clusters of two airport operational performance variables. Thus, the 3 in row 1 and column 1 of the 2x2 contingency

table means 3 airports are in common between the first cluster of airport operational variable 1 and the first cluster of airport operational variable 2.

**Table 7: Contingency Table - Sample**

		Airport operational variable 1		Total
		Cluster 1	Cluster 2	
Airport operational variable 2	Cluster 1	JFK LGA 3 ORD	0	3
	Cluster 2	SFO 2 BOS	BWI 1	3
Total		5	1	6

#### 4.5.2 Fisher’s Exact Test

Although visual representation of associations is helpful, it is more critical to further quantify the level of associations between two different clusters for this study. There are many statistical tests such as the Chi-Square Test, the G-Test, and Fisher’s Exact Test that determine level of association for categorical variables described by a Contingency Table. However, given that the number of entries in the Contingency Table is limited to 35 (only 35 OEP airports are studied), an assumption of a sufficiently large sample size for Chi-Square or G-Test does not hold. The Bernard Test is more accurate than Fisher’s Exact Test, but it is only applicable for 2x2 Contingency Tables. Thus, Fisher’s Exact Test is deemed as the most appropriate test to apply for our study [15-18]. However, the computation becomes burdensome when the size of the contingency table gets bigger for Fisher’s Exact Test. Thus, instead of calculating the exact *p*-values using Fisher’s exact test, we use Monte Carlo approximation to Fisher’s Exact Test with 50,000 samples (a sample of *p*-values

obtained from contingency tables can be found in Figure. 15). Note that the approximation error decreases proportionally to the square root of sample size. Thus, with a sufficiently large sample size, the error will be much smaller than the error introduced by using the Chi-Square Test. The  $p$ -values from the Monte Carlo approximation are used to indicate if one airport operational performance variable is related to the other airport operational performance variable. The null hypothesis of Fisher's exact test is that the two airport operational performance variables in the contingency table are independent[19]. So, if the  $p$ -value of the test is small, we can conclude either that the two variables are related, or that some extremely rare event occurred. Following standard statistical convention, we choose the former interpretation. Conversely, if the  $p$ -value is large, we can conclude that the two variables are not related.

	Hourly Utilization M1	Hourly Cap COV M1	Hourly Throughput M1	Hourly Throughput COV M1	Hourly CNX Ratio M1	Hourly ARR Delay M1	Hourly DEP Delay M1
Hourly Utilization M1	0.00	0.49	0.02	0.01	0.01	0.00	0.01
Hourly Cap COV M1	0.49	0.00	0.99	0.50	0.92	0.29	0.54
Hourly Throughput M1	0.02	0.99	0.00	0.03	0.04	0.16	0.39
Hourly Throughput COV M1	0.01	0.50	0.03	0.00	0.32	0.80	0.68
Hourly CNX Ratio M1	0.01	0.92	0.04	0.32	0.00	0.00	0.01
Hourly ARR Delay M1	0.00	0.29	0.16	0.80	0.00	0.00	0.00
Hourly DEP Delay M1	0.01	0.54	0.39	0.68	0.01	0.00	0.00

**Figure 15: Sample P values table should the pair-wise relationships between airport operational performance variables (Red=P value less or equal to 0.01, indicate strong relationship) <sup>2</sup>.**

<sup>2</sup> M1 = Similarity Metric 1, M2 = Similarity Metric 2

## Chapter 5: Quantitative Analysis

As mentioned in the last chapter, the pair-wise relationship between two airport performance variables is quantified by computing the  $p$ -value with respect to their contingency table. However, when the number of airport performance variables increases, the complexity of the  $p$ -values table in Figure 15 increases. After all the  $p$ -values are computed, it would be rather difficult to visualize the dependencies among the variables. The main purpose of the  $p$ -value is to determine if there is a statistical significance relationship between two variables. For this study, we chose 0.01 as the critical  $p$ -value to indicate the acceptable probability for a false positive. This critical  $p$ -value chosen is smaller than what might be considered a more traditional value of 0.05 to reduce the number of false positives of rejecting a null hypothesis when there is actually no relationship between two variables. Note that the table shown in Figure 15 contains approximately 400 pair-wise associations. Thus, if we chose the traditional  $p$ -value of 0.05, the expected number of false positives would be around 20 entries. However, if we reduce the critical  $p$ -value to 0.01, then the expected number of false positive is 4 entries. Thus, we have much more confidence in the trend in observed associations.

### 5.1 Design Structure Matrix (DSM)

Finally, when the table of Figure 15 is converted into binary variables using the critical  $p$ -value, it then resembles a Design Structure Matrix (DSM) [20, 21]. DSM is a classical project management tool that help engineering designers to manage complex systems. It is a square matrix representing relationships among

“components” in a system. Instead of using DSM to analyze a project/system, we employ the DSM concept to manage the  $p$ -value table so that the  $p$ -value results can be presented in a compact and concise way. Traditionally, engineering designers analyze DSM data by partitioning, tearing, and clustering result to extract more insights [22, 23]. In our case, DSM clustering is most appropriate because all we need from DSM is to highlight the important patterns in our results. In other words, by analyzing DSM with clustering algorithms, airport performance variables that are strongly interconnected are grouped together [24]. Related airport performance variables to the “target” variable can be quickly recognized from the clustered DSM graph as presented in section 5.2 below.

The DSM clustering algorithm used in this thesis was developed by Ronnie E. Thebeau [25]. The objective used to group the variable is to minimize the Coordination Cost. The basic idea is that the DSM clustering algorithm is trying to put as many variables in the “blue box” (see Figure 16) while preventing the size of the “blue box” from being too big (see Equation 12). The variables in the cluster (the “blue box”) are considered as having strong enough interconnection among them. Clustering algorithms for DSM is an active research area for the engineering design community. There are different ways to cluster the elements in the DSM. In this study, we employ one of the popular DSM clustering algorithms to cluster our results. Note that DSM clustering results in our study is intended to help us to visualize data. It does not change the results or conclusion of our analysis.

$$\min \{ TotalCost = \sum IntraClusterCost + \sum ExtraClusterCost \} \quad (13)$$

$$ExtraClusterCost = (DSM(j,k) + DSM(k,j)) * DSMSize^{powcc} \quad (14)$$

$$IntraClusterCost = (DSM(j,k) + DSM(k,j)) * ClusterSize(y)^{powcc} \quad (15)$$

where:

*TotalCost* = Coordination Cost

*IntraClusterCost* = Cost of interaction occurring within a cluster

*ExtraClusterCost* = Cost of interaction occurring outside of any cluster

*DSM(j,k), DSM(k,j)* = DSM interaction between elements *j* & *k*

*ClusterSize(y)* = Number of elements in the cluster *y*

*DSMSize* = Number of elements in the DSM

*powcc* = Penalizes the size of cluster

## 5.2 Results and Data Interpretation

There are a few different attainable results through the DSM in this analysis. If we use results from Figure 16, which are obtained by using all the data from 2002-2008, we can see the baseline pair-wise associations and isolated groups of performance variables. The other way to look at the results is to investigate the yearly results over time by plotting DSM per year. Note that we should keep the ordering the same so that we can visualize the trend. We want to look at the yearly results because certain patterns may reveal themselves by looking at the DSM graphs in chronological order, i.e. we may detect that an association between performance areas A and B did not exist until year 200x. Also, if we look at the quarterly DSM graphs, we may be able to detect seasonality; i.e. the association between performance area C and D only existed during winter terms. Finally, by combining yearly results into one DSM in Figure 24, the effect of uncertainty in our results can be reduced. As stated in Chapter 5, we should have approximately 4 false positives per DSM chart by using a *p*-value of 0.01. However, by “compressing” all the yearly results into one DSM as shown in Figure 24, we should be able to reduce these

uncertainty effects to a certain extent. Furthermore, we can also identify the significance of the relationships by counting the occasions of existence of the relationship (the green dots in Figure 17-23) over the years. The number of associations occurring over the years is represented by different colors and sizes of dots. Brown, red, yellow, orange, green, blue and violet represent 7, 6, 5, 4, 3, 2 and 1 time of occurrences respectively. The frequency of occurrences is also represented by the sizes of the dots. Associations represented by brown, red, and yellow are much less likely to be false positives. However, one of the possible drawbacks of analyzing the results in Figure 24 is that we must be careful to not to divide data too much to capture more the noise in the data than an actual trend (i.e. if we create DSM for each day, each DSM may not be as reliable as DSM created using yearly data). By comparing and contrasting Figure 16 and Figure 24, we have more confidence in our conclusion.

Here are some of the main findings from this analysis using yearly data:

- Each variable associates with itself.
- Variables with different similarity metrics are associated.
- Variables in blue boxes are related to different degrees.
- Arrival Delays has very strong relationship with departure delays.
- Number of runways is weakly related to capacity.
- Airport utilization is independent of airport hub structure.

It is essential to make sure each variable is self-associated. It is a way to self-validate the accuracy of the results because each variable must be strongly related to itself. Similarly, variables using different similarity metrics during the clustering

process should have relatively strong relationships to themselves because they are technically the same variables.

As shown in Figure 24, the number of runways has a very weak relationship (blue color) with capacity (both mean and COV). Therefore, it is very likely that number of runways cannot be used to capture airport capacity accurately. In the introduction section, we are using number of runways to estimate capacity in order to calculate airport utilization mainly because the capacity values for European airports are unavailable. Although number of runways is not the best way to capture capacity according to the results in this study, it is a way to look at airport capacity when capacity values are absent as they are still showing some level of association. Rather than using number of runways, using runway configurations may be a better representation for airport capacity when official capacity values are absent.

Furthermore, delay has some association with all of the variables in the blue box. More importantly, arrival delays generally have a strong relationship with departure delays. Also, we should consider using both arrival and departure utilization whenever we develop models to estimate or predict arrival delays. Even though there is a relationship between arrival delays and arrival utilization, it is stronger between arrival delays and overall (both arrival and departure) utilization (see Figure 24).

In this study, lack of association is as important as the existence of association because it points out some interesting phenomena. For example, hub information is not related to any other performance variables. Interestingly, even though hub airports have an image of being busy airports, utilization is seemingly independent of hub



structure. One possible explanation is that capacity at the hub airports is generally higher than at non-hub airports. Hence, even though hub airports have higher traffic compared to non-hub airports, utilizations of hub and non-hub airport are very similar. Other possible explanation of this result is that 1) the FAA's definition for hub may be insufficient to capture the nature of the airport hub structure, or 2) the hub list posted on the BTS website and used in some recent publications is out-of-date; therefore, there seems to be a necessity to develop a mathematical definition of hub that better relates to airport performance.

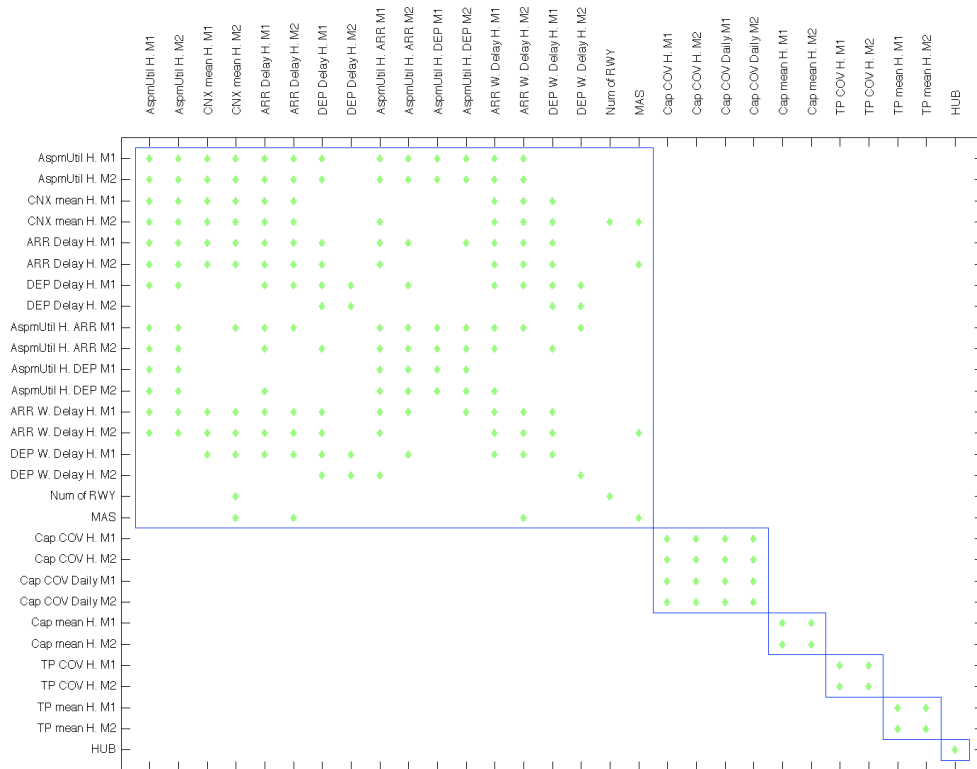
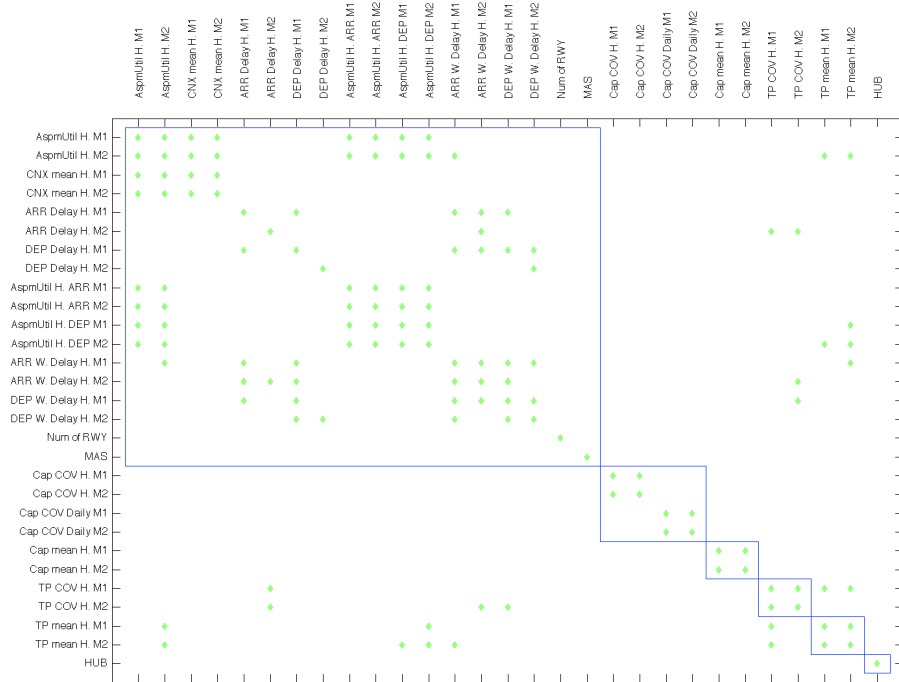
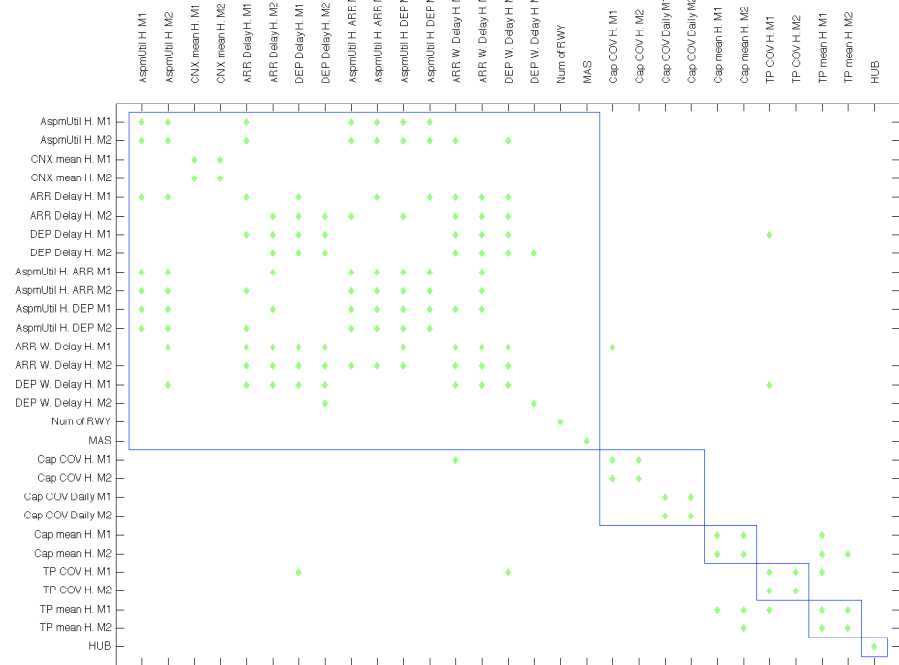


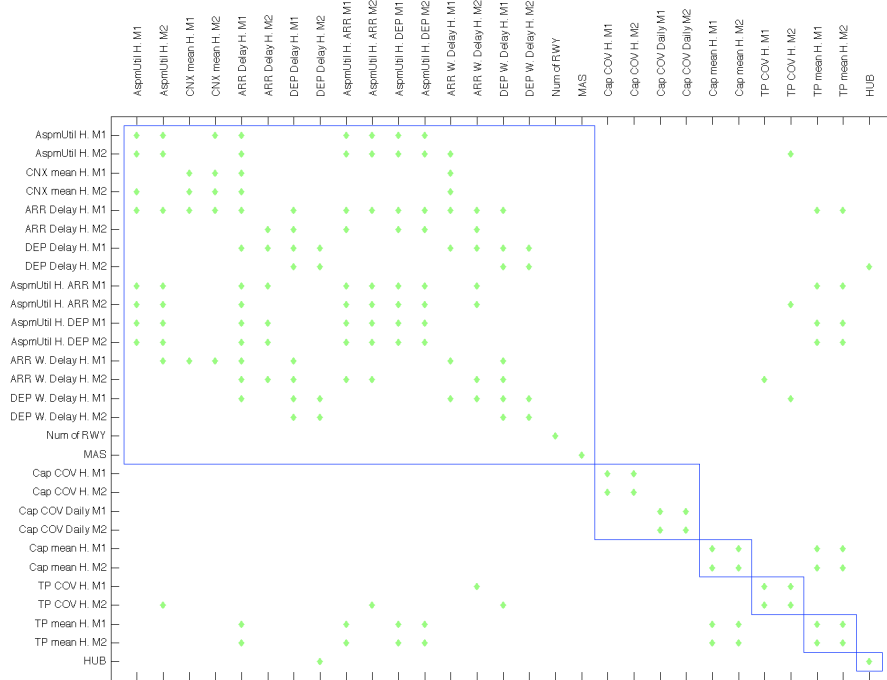
Figure 16: DSM representation of the results using data from 2002-2008



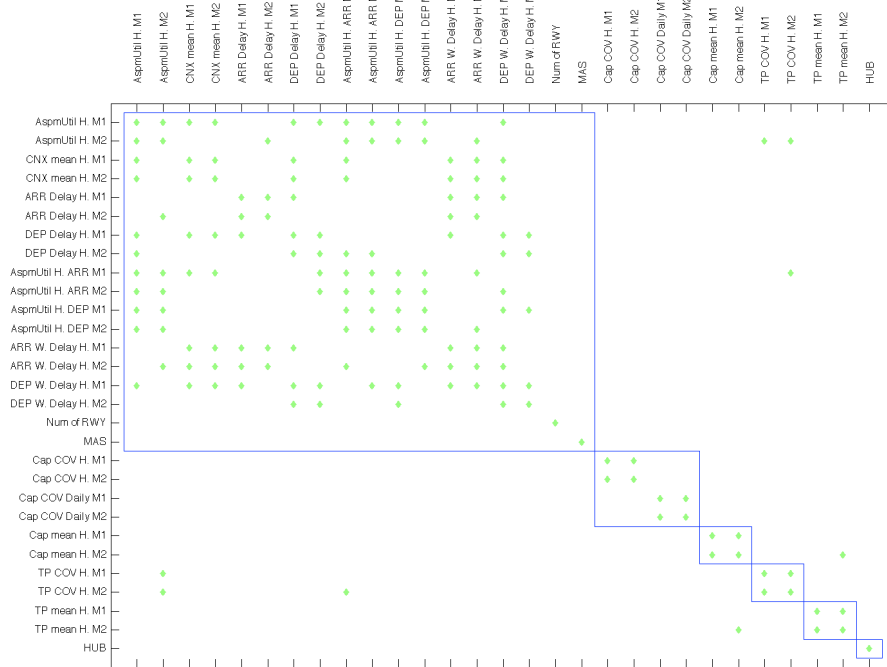
**Figure 17: DSM representation of the results using 2002 data**



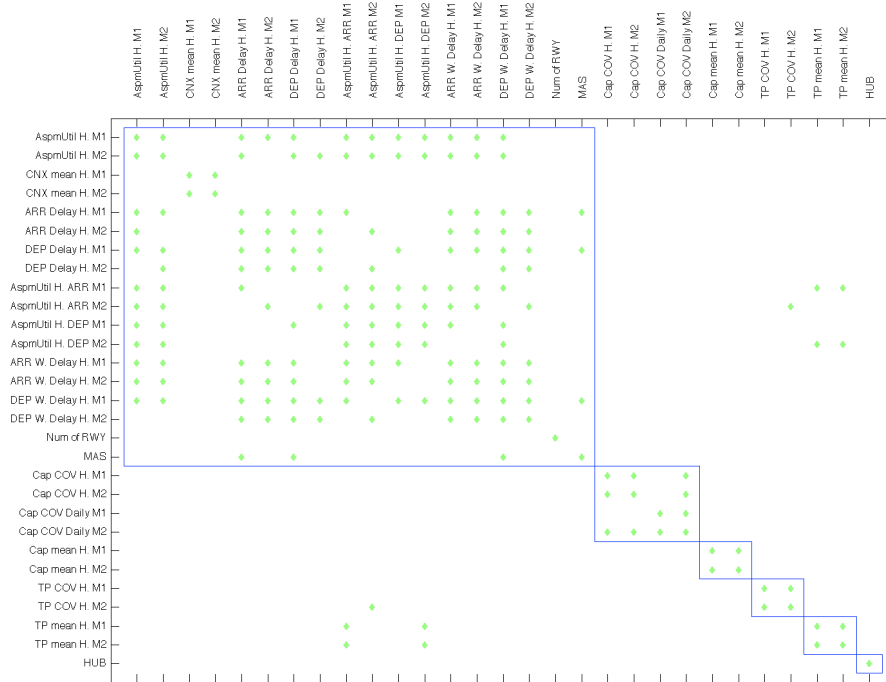
**Figure 18: DSM representation of the results using 2003 data**



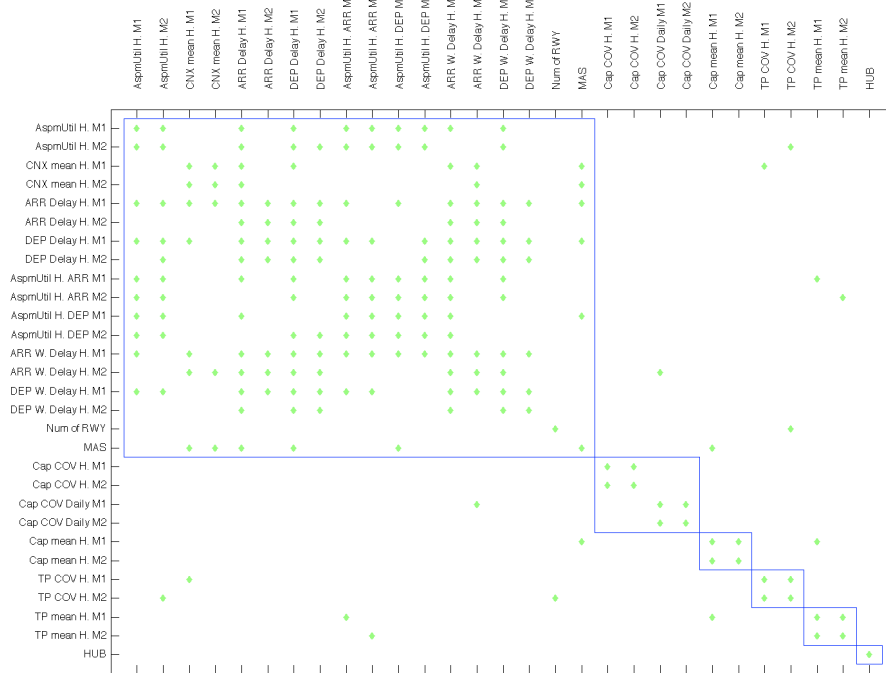
**Figure 19: DSM representation of the results using 2004 data**



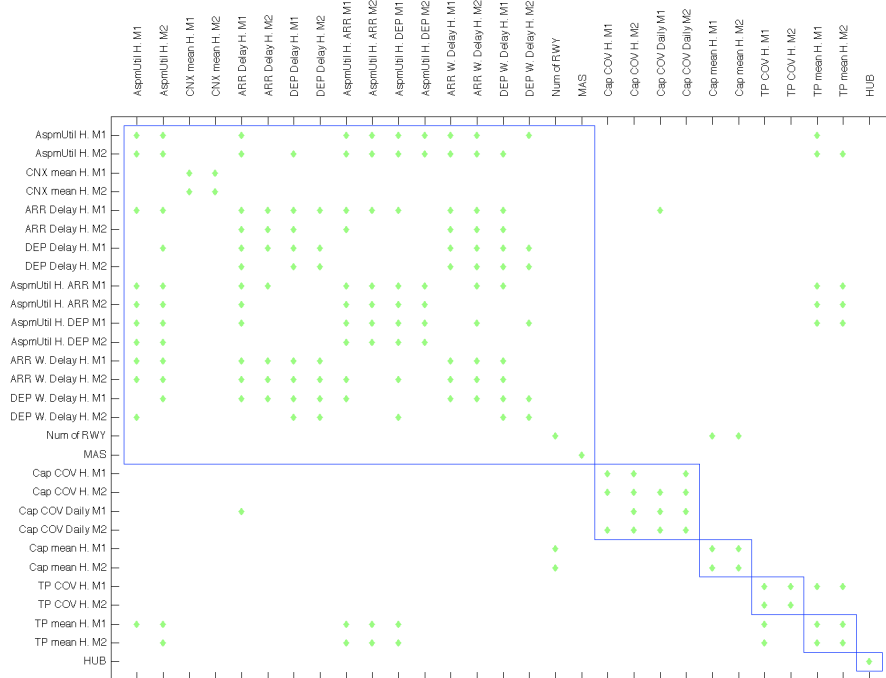
**Figure 20: DSM representation of the results using 2005 data**



**Figure 21: DSM representation of the results using 2006 data**



**Figure 22: DSM representation of the results using 2007 data**



**Figure 23: DSM representation of the results using 2008 data**

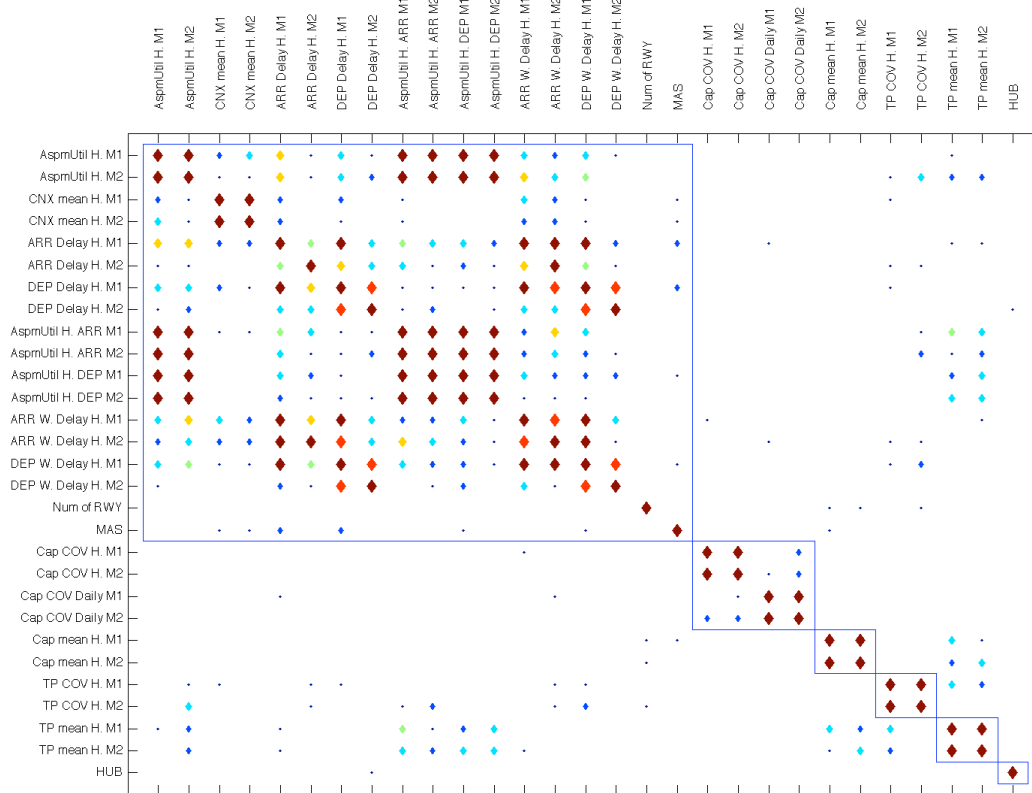


Figure 24: DSM representation of results using yearly data from 2002-2008<sup>3</sup>

### 5.3 Results Validation using Correlation analysis

The other way to validate the results from the proposed method is to compare them with the results generated from correlation analysis. To perform correlation analysis, we first need to convert the hourly airport data into scalars by taking an average of each variable. Then, we calculate the  $p$ -values for each pair of variables. Similarly to the proposed method, two variables are considered to be strongly related when the  $p$ -value is equal or less than 0.01. Finally, the correlation analysis results are displayed using DSM similar to the way we display the results for our study in Figure 24.

<sup>3</sup> The description of the variables in the figure can be found in the Appendix

The fundamental difference between correlation analysis and the proposed method is that we are investigating the pair-wise relationships between variables using the mean values rather than the distributions of the variables. Therefore, the results from the two methods should not be identical. However, they should have some similarity because some of the performances can actually be captured by the mean of the variables.

The results from the proposed method in Figure 24 and the results from correlation analysis in Figure 25 are comparable except the strength of the relationships is rather different for some variables. For example, both analyses show a relationship between utilization and throughput but the strength of the relationship is stronger based on the correlation results. Furthermore, the number of runways shows a much stronger relationship with capacity when using correlation analysis.

There are a few differences between the two results. For example, correlation analysis shows a strong relationship between the number of runways and airport throughput but our proposed method shows no relationship between them. Correlation analysis also indicates that there is no relationship between delays and throughput; however, the proposed method shows a weak relationship between them.

Due to the similarities of Figures 24 and 25, correlation analysis does validate the results of our study. However, the results from the correlation analysis may not be as reliable as our proposed method because it is very likely that critical information is removed when we condense the data into scalar form for some of the airport performance variables while doing correlation analysis. Theoretically, averaging is a form of filtering to remove variations. However, it has to be done carefully to prevent

us from removing important patterns in data such as patterns in peak and off-peak hours. Traffic during peak hours and off peak hours behave very differently and, therefore, we should not averaging peak traffic with off peak traffic. The other advantage of using our proposed method is that it can analyze pair-wise relationships for scalar variables as well as vector variables.

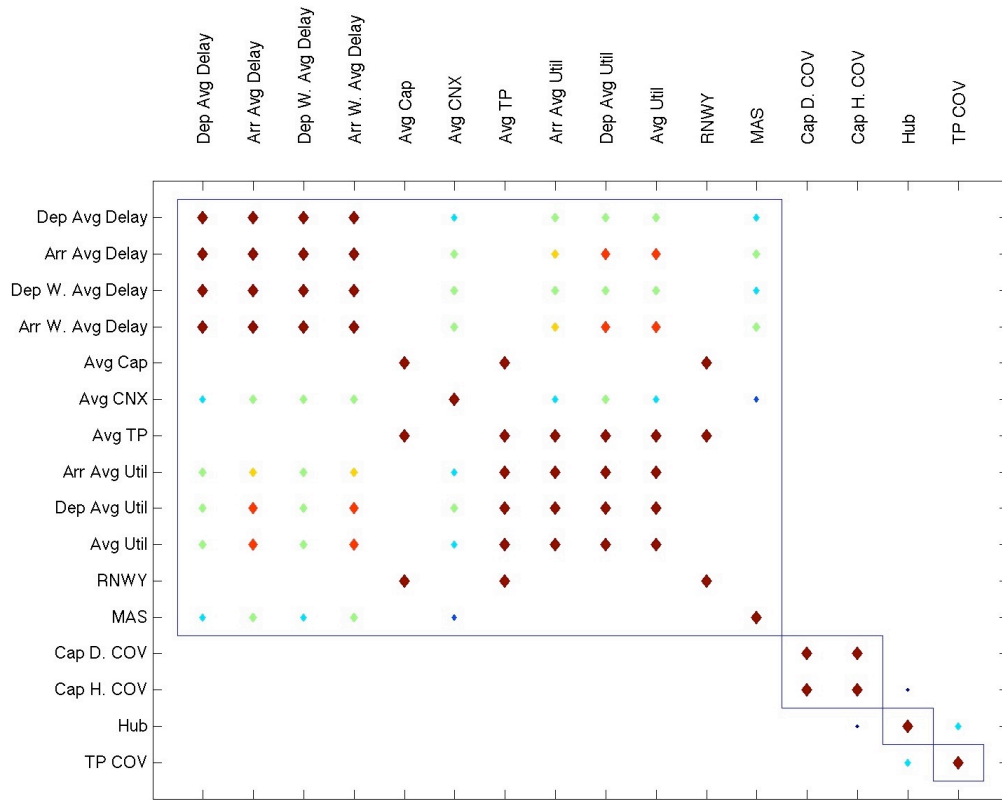


Figure 25: DSM representation of correlation analysis results using yearly data from 2002-2008<sup>4</sup>

<sup>4</sup> The description of the variables in the figure can be found in the Appendix.



## Chapter 6: Conclusions

In this study, we point out that a better understanding of the inter-relationships between airport operational performance variables is important in the process of analyzing operational performance at airports. Therefore, we propose a method to investigate the pair-wise relationships between performance variables, to take advantages of the extensive amount of data stored in various FAA and European air traffic performance databases. The proposed method is illustrated by using some of the popular airport operational performance variables and the results suggested that arrival delays, weighted arrivals delays, departure delays, weighted departure delays, utilization, arrival utilization, departure utilization, MAS, number of runways and cancellation ratio are related to different degrees.

The proposed method consists of the following steps:

- 1) Cluster airports using each airport performance variable separately.
- 2) For each pair of performance variables, create contingency tables of the two airport performance variables. Then, compute the  $p$ -value associated with each contingency table.
- 3) Use these  $p$ -values to construct the Design Structure Matrix.
- 4) Cluster the DSM to arrange the most associated variables into the same groups.

By representing the results in the DSM, relationships between airport performance variables can be visually understood. Also, by using more complete and accurate data and definitions of airport performance variables, the results of the proposed method can be used to develop more accurate models to help policy makers

decide how resources should be allocated. It is very important to have the basic understanding of what we model before constructing a model to answer certain air traffic related questions, as the modeling process is very time-consuming and costly. Furthermore, the results of the proposed method can also be useful during the model validation process. With the basic understanding of the relationships between variables in the model, 1) extra variables can be eliminated to prevent over-fitting when building models using historical data and 2) crucial variables will not be ignored to prevent under-fitting of the model.

The method proposed in this thesis can help to answer a variety of air traffic questions accurately after further analysis. Questions such as the following can be answered precisely by developing models using the findings from the proposed method:

- Should European airports increase the number of allocated landing slots?
- Why do we see different delay patterns in Europe and U.S?
- Should we limit the maximum capacity (i.e., impose operational caps) to alleviate the congestions at some of the U.S airports, and if so, which airport(s) should be selected?
- What performance gains should be expected if we build a new runway at airport X?

The proposed method can also be used to provide assistance for the FAA to select candidate models during the funding process, for example to determine which

model reflects reality the most, to evaluate the usability of the models and the reliability of the modeling results.

In closing, the methodology suggested in this thesis intends to help researchers or aviation analysts to understand airport operations better, which will eventually help decision makers in allocating resources in the right ways.

## Chapter 7: Future Work

There are many possible extensions to this thesis work. First, notice that data quality is crucial for this methodology. Therefore, a better and larger set of data should be obtained from both Europe and the U.S. Furthermore, we have demonstrated that an airport performance definition is crucial for applying this technique. Thus, one could determine more appropriate airport performance variables by exploring the different definitions and derivations of airport performance variables. We should also include some other quantifiable data such as weather into the analysis. Finally, we should then build models to analyze some of the air traffic performance questions such as why there are different delay patterns in Europe and the U.S by using the key variables selected from the proposed methodologies in this thesis.

## Appendices

Airport Operational Performance Variable 1	Airport Operational Performance Variable	Description
AspmUtil_hourly_M1	AspmUtil H.M1	Hourly Airport Utilization using ASPM data and similarity metric 1
AspmUtil_hourly_M2	AspmUtil H.M2	Hourly Airport Utilization using ASPM data and similarity metric 2
Cap_COV_hourly_M1	Cap COV H.M1	Hourly Capacity coefficient of variation (COV) using similarity metric 1
Cap_COV_hourly_M2	Cap COV H.M2	Hourly Capacity coefficient of variation (COV) using similarity metric 2
Throughput mean hourly_M1	TP mean H.M1	Hourly Throughput using similarity metric 1
Throughput mean hourly_M2	TP mean H.M2	Hourly Throughput using similarity metric 2
Throughput COV hourly_M1	TP COV H.M1	Hourly Throughput coefficient of variation (COV) using similarity metric 1
Throughput COV hourly_M2	TP COV H.M2	Hourly Throughput coefficient of variation (COV) using similarity metric 2
CNX_mean_hourly_M1	CNX mean H.M1	Hourly Cancellation Ratio using similarity metric 1
CNX_mean_hourly_M2	CNX mean H.M2	Hourly Cancellation Ratio using similarity metric 2
ARR_Delay_hourly_M1	ARR Delay H.M1	Hourly Arrival Delay using similarity metric 1
ARR_Delay_hourly_M2	ARR Delay H.M2	Hourly Arrival Delay using similarity metric 2
Dep_Delay_hourly_M1	Dep Delay H.M1	Hourly Departure Delay using similarity metric 1
Dep_Delay_hourly_M2	Dep Delay H.M2	Hourly Departure Delay using similarity metric 2
AspmUtil_hourly_ARR_M1	AspmUtil H.ARR M1	Hourly Arrival Utilization using ASPM data and similarity metric 1
AspmUtil_hourly_ARR_M2	AspmUtil H.ARR M2	Hourly Arrival Utilization using ASPM data and similarity metric 2
AspmUtil_hourly_DEP_M1	AspmUtil H.DEP M1	Hourly Departure Utilization using ASPM data and similarity metric 1
AspmUtil_hourly_DEP_M2	AspmUtil H.DEP M2	Hourly Departure Utilization using ASPM data and similarity metric 2
Cap_mean_hourly_M1	Cap_mean H.M1	Hourly Capacity using similarity metric 1
Cap_mean_hourly_M2	Cap_mean H.M2	Hourly Capacity using similarity metric 2
Cap_COV_Daily_M1	Cap COV Daily M1	Daily Capacity coefficient of variation (COV) using similarity metric 1
Cap_COV_Daily_M2	Cap COV Daily M2	Daily Capacity coefficient of variation (COV) using similarity metric 2
ARR_Weighted_Delay_hourly_M1	ARR W.Delay H.M1	Weighted hourly Arrival Delay using similarity metric 1
ARR_Weighted_Delay_hourly_M2	ARR W.Delay H.M2	Weighted hourly Arrival Delay using similarity metric 2
Dep_Weighted_Delay_hourly_M1	Dep W.Delay H.M1	Weighted hourly Departure Delay using similarity metric 1
Dep_Weighted_Delay_hourly_M2	Dep W.Delay H.M2	Weighted hourly Departure Delay using similarity metric 2
	Hub	Hub size
	MAS	Multiple Airport System
	Num Of RWY	Number of Runways

**Figure 26: Description of variables**

**Table 8: ASPM carriers**

	<b>Air Carriers</b>
1	Air Canada (ACA)
2	Airtran Airways TRS*
3	Alaska Airlines (ASA)*
4	Aloha Airlines (AAH)*
5	American Airlines (AAL)*
6	American Eagle (EGF)*
7	America West (AWE)*
8	ATA Airlines (AMT)*
9	Atlantic Coast (BLR)*
10	Atlantic Southeast Airlines (ASQ)*
11	Atlantic Southeast Airlines (CAA)*
12	Comair (COM)*
13	Continental Airlines (COA)*
14	Delta Air Lines (DAL)*
15	ExpressJet Airlines (BTA)*
16	FedEx (FDX)
17	Frontier Airlines FFT*
18	Hawaiian Airlines HAL*
19	Independence Air IDE*
20	Jetblue Airways JBU*
21	Mesa Airlines (ASH)*
22	Northwest Airlines NWA*
23	Pinnacle Airlines (FLG)
24	Skywest Airlines SKW*
25	Southwest Airlines SWA*
26	TWA (TWA)*
27	United Airlines (UAL)*
28	United Parcel Service (UPS)
29	US Airways (USA)*

Note: Although some of these carriers may no longer be in operation, ASPM has tracked operations for them since January 2000.

\* Denotes an ASQP Carrier

Table 9: ASPM airports

	<b>Airport ID</b>	<b>Airport</b>
1	ABQ	Albuquerque Intl Sunport
2	ANC	Ted Stevens Anchorage Intl
3	ATL	Hartsfield-Jackson Atlanta Intl*
4	AUS	Austin-Bergstrom Intl
5	BDL	Bradley Intl
6	BHM	Birmingham Intl
7	BNA	Nashville Intl
8	BOS	Boston Logan Intl*
9	BUF	Buffalo Niagara Intl
10	BUR	Bob Hope (Burbank/Glendale/Pasadena)
11	BWI	Baltimore/Washington Intl*
12	CLE	Cleveland Hopkins Intl*
13	CLT	Charlotte Douglas Intl*
14	CVG	Cincinnati/Northern Kentucky Intl*
15	DAL	Dallas Love Field
16	DAY	Dayton Intl
17	DCA	Ronald Reagan Washington National*
18	DEN	Denver Intl*
19	DFW	Dallas/Fort Worth Intl*
20	DTW	Detroit Metropolitan Wayne County*
21	EWR	Newark Liberty Intl*
22	FLL	Fort Lauderdale/Hollywood Intl*
23	GYG	Gary Chicago Intl
24	HNL	Honolulu Intl*
25	HOU	Houston Hobby
26	HPN	Westchester County
27	IAD	Washington Dulles Intl*
28	IAH	George Bush Houston Intercontinental*
29	IND	Indianapolis Intl
30	ISP	Long Island Mac Arthur
31	JAX	Jacksonville Intl
32	JFK	New York John F. Kennedy Intl*
33	LAS	Las Vegas McCarran Intl*
34	LAX	Los Angeles Intl*
35	LGA	New York LaGuardia*
36	LGB	Long Beach
37	MCI	Kansas City Intl
38	MCO	Orlando Intl*
39	MDW	Chicago Midway*
40	MEM	Memphis Intl*
41	MHT	Manchester
42	MIA	Miami Intl*
43	MKE	Milwaukee Gnl Mitchell International
44	MSP	Minneapolis/St. Paul Intl*
45	MSY	Louis Armstrong New Orleans Intl
46	OAK	Oakland Intl
47	OGG	Kahului
48	OMA	Omaha Eppley Airfield
49	ONT	Ontario Intl
50	ORD	Chicago O'Hare Intl*
51	OXR	Oxnard
52	PBI	Palm Beach Intl
53	PDX	Portland Intl*
54	PHL	Philadelphia Intl*
55	PHX	Phoenix Sky Harbor Intl*
56	PIT	Pittsburgh Intl*
57	PSP	Palm Springs International
58	PVD	Providence Francis Green State
59	RDU	Raleigh/Durham Intl
60	RFD	Greater Rockford
61	RSW	Southwest Florida Intl
62	SAN	San Diego Intl*
63	SAT	San Antonio Intl
64	SDF	Louisville Intl
65	SEA	Seattle/Tacoma Intl*
66	SFO	San Francisco Intl*
67	SJC	Norman Mineta San Jose Intl
68	SJU	San Juan Luis Munoz Intl
69	SLC	Salt Lake City Intl*
70	SMF	Sacramento International Airport
71	SNA	John Wayne Airport-Orange County
72	STL	Lambert Saint Louis Intl*
73	SWF	Stewart Intl
74	TEB	Teterboro
75	TPA	Tampa Intl*
76	TUS	Tucson Intl
77	VNY	Van Nuys

\* Denotes an 35 OEP airport

## References

- [1] Sarkis, J., 2000, "An analysis of the operational efficiency of major airports in the United States " *Journal of Operations Management*, 18, pp. 335-351.
- [2] Knorr, D., Rose, M., and Bonn, J., 2009, "US/ Europe comparison of ATM-related operational performance " Napa, CA.
- [3] Neufville, R. D., and Odoni, A. R., 2002, *Airport Systems: Planning, Design, and Management*, McGraw-Hill Professional Publishing
- [4] John Gulding, D. K., Marc Rose, James Bonn, Philippe Enaud, Holger Hegenoefer, 2009, "US/ Europe comparison of ATM-related operational performance " Napa, California, USA
- [5] Wine, C., 2005, "Presentation of System Airport Efficiency Rate (SAER) and Terminal Arrival Efficiency Rate (TAER)."
- [6] Federal Aviation Administration, *Airport Categories*, October 23, 2009, [http://www.faa.gov/airports/planning\\_capacity/passenger\\_allcargo\\_stats/categories/](http://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/categories/)
- [7] Federal Aviation Administration, *ASPM System Overview*, 11/5/2009, [http://aspmhelp.faa.gov/index.php/ASPM\\_System\\_Overview](http://aspmhelp.faa.gov/index.php/ASPM_System_Overview)
- [8] Bureau of Transportation Statistics, *Airline On-Time Performance Data*, 11/5/2009, [http://www.transtats.bts.gov/DatabaseInfo.asp?DB\\_ID=120&DB\\_URL=Mode\\_ID=1&Mode\\_Desc=Aviation&Subject\\_ID2=0](http://www.transtats.bts.gov/DatabaseInfo.asp?DB_ID=120&DB_URL=Mode_ID=1&Mode_Desc=Aviation&Subject_ID2=0)
- [9] 2009, *Enhanced Traffic Management System (ETMS)*, <http://hf.tc.faa.gov/projects/etms.htm>
- [10] *Enhanced Traffic Management System Counts (ETMSC)*, <http://aspmhelp.faa.gov/index.php/ETMSC>
- [11] AirNav, LLC, *AirNav*, 11/5/2009, <http://airnav.com/>
- [12] *Avigation Networks, Inc, World Aero Data* 11/5/2009, <http://worldaerodata.com/>
- [13] Jung, Y., Park, H., Du, D.-Z., and Drake, B. L., 2003, "A Decision Criterion for the Optimal Number of Clusters in Hierarchical Clustering," *Journal of Global Optimization*, 25, pp. 91-111.
- [14] Joseph F. Hair, J., Anderson, R. E., Tatham, R. L., and Black, W. C., 1998, *Multivariate Data Analysis*, Prentice Hall.
- [15] Mato, A. S., and Andre as, A. M., 1997, "Simplifying the calculation of the P-value for Barnard's test and its derivatives " *Statistics and Computing*, 7, pp. 137-143.
- [16] McDonald, J., 2009, *Handbook of Biological Statistics* Sparky House Publishing.
- [17] McDonald, J. H., 2009, *Handbook of Biological Statistics* Sparky House Publishing.
- [18] 2009, *Chi-square test of independence*, <http://udel.edu/~mcdonald/statchiind.html>
- [19] 1996, *Statistics to use*, <http://www.physics.csbsju.edu/stats/exact.html>
- [20] Steward, D. V., 1981, "The Design Structure System: A Method for Managing the Design of Complex System," *IEEE Transactions on Engineering Management*, 28, pp. 71-74.
- [21] Steward, D. V., 1981, *Systems Analysis and Management: Structure, Strategy and Design.*, Petrocelli Books, New York, NY.
- [22] Gebala, D. A., and Eppinger, S. D., 1991, "Methods for Analyzing Design Procedures," *ASME*, Miami, FL.

[23] Avnet, M. S., 2009, "Socio-cognitive Analysis of Engineering Systems Design: Shared knowledge, process, and product," Doctor of Philosophy in Engineering Systems Massachusetts Institute of Technology.

[24] Institute of Product Development Technische Universität München, 11/10/2009, <http://129.187.108.94/dsmweb/en/understand-dsm/technical-dsm-tutorial0/clustering.html>

[25] Thebeau, R. E., 2001, "Knowledge Management of System Interfaces and Interactions for Product Development Process," M.S Massachusetts Institute of technology.