International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Assessing Aircraft Timeliness Variations By Major Airlines: Passenger Travel Practice In Uganda

Wesonga Ronald[a]*, Nabugoomu Fabian[b], Masimbi Brian[c]

[a,c]School of Statistics and Planning, Makerere University, Kampala, Uganda

[b]Kyambogo University, Kampala, Uganda

[a]E-mail: wesonga@wesonga.com

[b]E-mail: fnabugoomu@yahoo.com

[c]E-mail: bmasimbi@yahoo.com

Abstract

Flight delays do not only affect passenger satisfaction but also carry along costly consequences to airlines. The overall objective of the study was to assess aircraft timeliness variations by major airlines so as to determine passenger travel practice in Uganda. The study hypotheses were tested using a two-way ANOVA F-test and further measures of associations.

The study found out that the number of schedules of each airline per day had a positive effect on the delay duration, whereby an additional schedule increased the average delay by a proportion of 11%. Whereas the day of the week $F(16, 1129) = 1.36$, $p > 0.01$ had no significant difference in the delays amongst the airlines, the month of the year $F(33, 1107) = 1.88$, $p < 0.001$ showed a significant difference. However, the total variance of the delays was attributed to the airline (29%). It was also demonstrated from the analysis that Eagle Air (EA), Kenya Airways (KA) and South African Airways (SAA) experienced more delays than the British Airways (BAW) by 33%, 62% and 55% respectively. Other than Wednesday, flights were delayed more on all the days of the week and less delayed in the months of October and November than in June by 26% and 3% respectively.

On Saturdays and Sundays, flights were found to have longer periods of delay ($p<0.05$) that averaged 14 and 13 minutes respectively. The flights in January and March had longer delays (15 and 14 minutes) than that recorded in the other months. Therefore, it can be concluded that the passengers who use BAW are less likely to delay than the other (EA, KA and SAA) airlines and travelling in the months of October and November is highly recommended. Given that airline delay is positively correlated with the number of scheduled flights, a policy framework could be developed to optimise schedules and airline delays during departure at the airport. The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them.

Keywords: Airport management; departure delay; analysis of variance; month; day

----

* Corresponding author.

E-mail address: wesonga@wesonga.com.

1. Introduction

Air traffic has had a positive trend since the year 2000.This rapid increase in air traffic flow arises from increased demand for airport and airspace resources. Suffice to note that International passenger numbers sprung from 118,000 in 1991 to over one million by 2011; a factor that has led to more congestion and delays [1]. Flight delays do not only affect passenger satisfaction but also bring along costly consequences to airlines [2] that are said to have been dire if the country had a national airline.

According to[3], in the year 2007, sixteen international airlines had scheduled operations to and from Entebbe International Airport, serving fourteen different destinations. Uganda's geographical location in the heart of Africa gives Entebbe International Airport (EIA) greater advantage for hub and spoke operations in the Eastern and Southern African region. Consequently, EIA has developed as Uganda's international gateway with its traffic growing from a paltry 27000 passengers in 1962 to 1.2 Million at the turn of Uganda's 50th Independence anniversary [1]. The growth in traffic has inadvertently caused significant delays in air traffic flow.

At European airports, four significant variables were found to explain delays; market concentration, slot coordination, hub airports and hub airlines [4]. It was established that in order to appropriately understand aircraft delays at the airport, both aviation and meteorological variables should be investigated [5]. Delay could be categorized into primary and secondary delays [6]. A primary delay cause may be defined as delay that affects the initiation of the flight. This delay is unaffected by any earlier or accumulated delay. However, a reactionary delay is accumulated. This is a delay imposed as a consequence of the unavailability of aircraft, crew or load due inbound by the scheduled departure time because of a disruption earlier in the day [7]. The cause of the earlier disruption could be itself either a primary delay at the start of the previous flight, or a reactionary delay arising from an even earlier incident. It has proved to be complex and difficult for airlines to record successfully the precise origins of reactionary delay, and these difficulties have outweighed the benefits. For this reason, and the logical extension, which states that there will have been (in 99% of cases) a primary cause, the recording and collection of primary delay and its causes has been a feasible and effective way to monitor operational performance.

Analysis of delay can be summarized into five categories: Regression and related methods, Time series analysis, Bayesian Networks model analysis, Cluster and classification analysis and Simulation. It was further noted regression includes methods for using observations to predict or explain the delays. The methods include linear regression, neural networks and related methods. In time series analysis, the trend analysis, spectral analysis and Markov Chain analysis are introduced. Linear classification and cluster analysis are under classification. [8]used regression (Recursive OLS) in the analysis of arrival delays given different causes of the delays.

CAA in coordination with Airline Operation Centres manages the arrival and departure flow of aircraft at EIA based on their rates. Inefficiencies in the traffic flow occur when non-traffic flow delays (e.g. carrier) are super imposed on the traffic flow delays. The researchers have correlated the causes of non-traffic flow delays at airports with sets of casual factors and have created models to predict aggregate delays at airports on the time scale of a day. In order for this study to be consistent with the way traffic is managed, a model of casual factors of delays would provide the analytical bases for improving the efficiency of Traffic Flow Management at Entebbe International Airport.

The overall objective of the study was to assess aircraft timeliness variations by major airlines so as to determine passenger travel practice in Uganda [9]. And the specific objectives of the study were; to determine the effect of the number of schedules for each airline per day on the aircraft delays, to determine whether different days of the week impact on average airline delay depend on the type of the airline and also to determine whether different months' impact on average airline delay depend on the type of the airline. The research hypotheses of the study were:

$H_{01:}$ the number of schedules for each airline per day has no effect on the aircraft delays.

H_{o2}: the different days of the week's impact on average airline delay does not depend on the flight airlines

H_{o3}: the different month's impact on average airline delay does not depend on the flight airlines.

The study covered Entebbe International Airport (EIA) which handles all the International flights that come into Uganda and it has one runway which handles both Arrivals and Departures. The airlines selected for this study were Kenya Airways, South Africa Airways, Eagle Airline and British Airways. The time scope of the study was January to December, 2008.

The significance of the study was to establish a model that would help to predict airline delays. It also investigated the cause of delays at airports and also helped in knowing the performance of aircrafts from the four airlines. The information generated here would be important when deriving solutions to abate airline departure delay and their associated characteristics [10].

2. Materials and Methods

2.1 Data management process

Aviation data was obtained in a Microsoft excel format with different sheets each storing daily data for a specific month of the year 2008. The data covered twelve months of year 2008 The dataset had 8 attributes namely Date, Operator (Airline/Carrier), Type of Aircraft, Nationality (Origin), From/To (where the Aircraft is coming from and its destination), category (international/Local), Expected Time of arrival/Departure (ETA/D) and Actual time of Arrival/Departure.

A sample of four airlines was taken; the data for four airlines considered being busy that is, with the highest number of inflows and outflows used included; Eagle Air, Kenya Airways, South African Airways and British Airways. To achieve the stated objectives and hypotheses in this research, four fields which deemed relevant were extracted from the dataset. These included Date, Operator (Airline), Scheduled Time of Departure/Arrival (ETA/D) Actual Time of Departure/Arrival (ATA/D).

Four derived data fields were added to each record namely Month, Day of the week, Delays (minutes) and Schedule per day (Number of schedules per airline per day/Count). The derived data fields represent values that are computed from the primary data.

The formula below was used in the computation of the delays for each record:

$$Delay_i(minutes) = \frac{ATA}{D}_i - \frac{ETA}{D}_i \qquad (1)$$

Where i is a real number (1, 2, 3…)

For the Schedules per airline per day, a count was done

It was assumed in this study that any earlier departures and arrivals were not delayed thus assigned a value zero. Furthermore, a record with missing data on scheduled time of departure/arrival (ETD/A) and actual time of departure/arrival (ATD/A) or both were eliminated from the dataset because the missing data could not easily be estimated.

Since in most scenarios, the airlines had more than one schedule a day at Entebbe International Airport, the data was further aggregated into daily averages for each airline per day [11].

2.2 Normality Test for the delays

A normality test was done on the dependent variable using the Shapiro-Wilk test. Findings from the test (W=0.63388, P-value = 0.0000, N = 504) implied that the delays are not normally distributed. As is a requirement in regression modelling that the independent variable should be normally distributed, so a natural log transformation was done on the delay variable and test results (W=0.98754, P-value = 0.00055, N = 462) implied that the log transformations are normally distributed.

Three other variables were computed as numerical representatives of the variables day of the week, Month and Airlines

a) Day (1-Monday, 2-Tuesday, 3-Wednsday, 4-Thursday, 5-Friday, 6-Saturday, 7-Sunday)
b) Airline1(1-Eagle Air, 2-British Airways, 3-Kenya Airways, 4-SouthAfrican Airways)
c) Month(1-January, 2- February, 3-March, 4-April, 5-May, 6-June, 7-July, 8-August, 9-September, 10-October, 11-November, 12-December)

The final dataset used in this study had the following variables namely; Dates: day of the week, date, month and year, Arrival and departure times: actual and scheduled, Airline: Eagle air, Kenya Airways, South African Airways and British Airways: Average Delays, lndelays (natural logarithms of delays), Month, Day and Airline1, Schedule per Day.

2.3 Theory and Calculation

First and foremost a regression analysis was performed for the delays and the number of schedules for the airlines. Secondly a plot analysis for the categorical variables in order to find out if they led to delay. The first plot was delay (minutes) by the day of the week, this plot established if the day of the week affected the airline delay. The second plot was the departure delay (minutes) by the Airline, it showed out how long each plane delayed on average.

A two-way ANOVA was used to test the hypotheses for this study. It can be noted that a two-way ANOVA focused on group means because it was an inferential technique, any two way ANOVA was actually concerned with the set of μ values that correspond to the sample means that are computed. Statistically significant findings deal with means i.e. statistical assumption needed to be tested and the research questions dictated whether planned or post hoc comparisons are used in conjunction with the two-way ANOVA. A factorial ANOVA was used to address research hypotheses that focused on the difference in the means of one dependent variable when there are two or more independent variables. Further a measure of Association ω^2 (omega squared) was calculated which indicated the promotion of variance in the dependent variable that was accounted for by the levels of independent variable. The use of ω^2 can be extended to two ways ANOVA with the general formula;

$$\omega^2 = \frac{SS_{effect} - (df_{effect})(MS_w)}{SS_T + MS_w} \qquad (2)$$

Post-hoc tests were performed where the main effects were found for between-subjects factors with three or more levels while post hoc comparisons were performed with the Dun-Sidak test (family wise $\alpha = 0.05$). This test does not require the overall F-test for groups to be significant as they control the family wise error rate independently and test different hypotheses from the overall ANOVA with different power (Howell, 1997, p. 351). Where significant interactions were found following factorial analysis of variance, simple effects of a priori interest were calculated by one-way ANOVA and tested by hand against the pooled error term ($F = MS_{factor}/MS_{pooled\ error}$; critical values of F based on df_{factor} and $df_{pooled\ error}$). Multiple comparisons for simple effects were performed as described above but using the pooled error term.

However, where significant interactions were found following repeated measures' analysis, a pooled error term was used to test between-subjects simple effects of a priori interest, but separate error terms, that is one-way ANOVA were used for within-subjects' factors as sphericity corrections are inadequate if a pooled error term is used [12].

3. Results

3.1 Delays and the Number of Schedules Regression

In this case we assume that the number of schedules for each airline per day is a good predictor variable of the delays. Table 1 shows a simple least squares regression predicting delays by the number of schedules.

Table 1 Delay and number of schedules

Term	Estimate	Std. Err.	t Ratio	p>[t]
Schedule	0.058	0.011	5.23	0.000
Cons	1.745	0.056	31.33	0.000

From the results, the line fit yielded a slope of 0.06, meaning that for every unit increase in the number of schedules, the delays increase by 0.06 minutes and further it was noted that the number of schedules has a positive effect on the delays (minutes). But with the R squared at 0.0232, it is clear that the number of schedules does not account for most of the variance in the delays.

It seems therefore, that some other factors besides schedules cause variance in the delays. We then examined how day of the week affects duration of delay. The effect of the day on the delays was examined as shown in Figure 1.

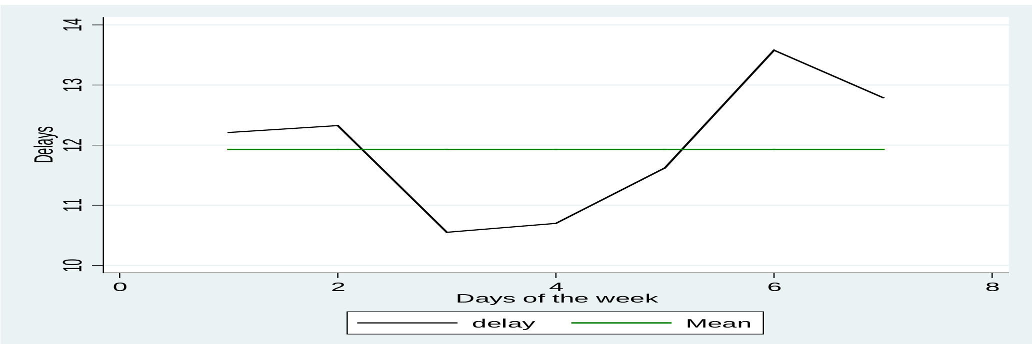

Figure 1: Delay duration in minutes by day of the week

Figure 1 shows that there was a stronger relationship between Day and the duration of delay. It can be seen that flights on Monday, Tuesday, Saturday and Sunday arrive later than flights on the others days. Having sampled a full year (2008) of flights, the relationship between day and delays should be fairly sound, as it is unlikely that certain external sources of variation hit the same days of the week for a full year. The effect of the airline operator on the delays was examined as below as shown in Figure 2.

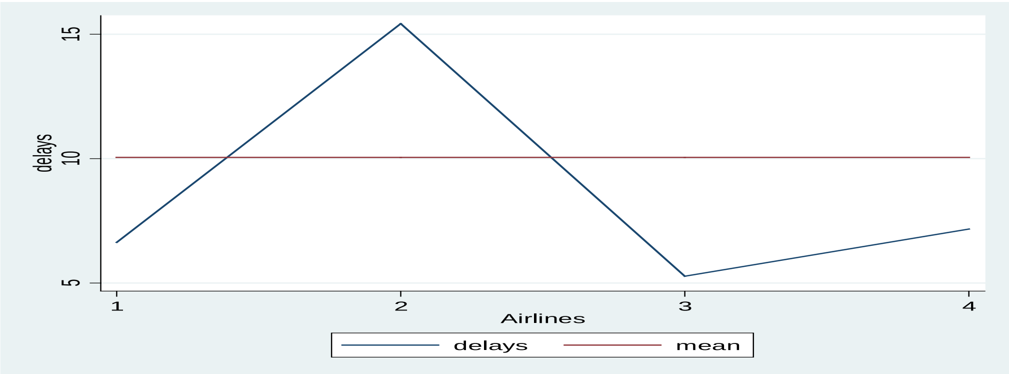

Figure 2: Delay duration in minutes by Airline

It was noted Eagle Air appear to be approximately 6 minutes early on average while British Air, Kenya Airways and South African Airways appear to be 5 minutes early on average.

The effect of the Month on the duration of delay in minutes was examined as in Figure 3.

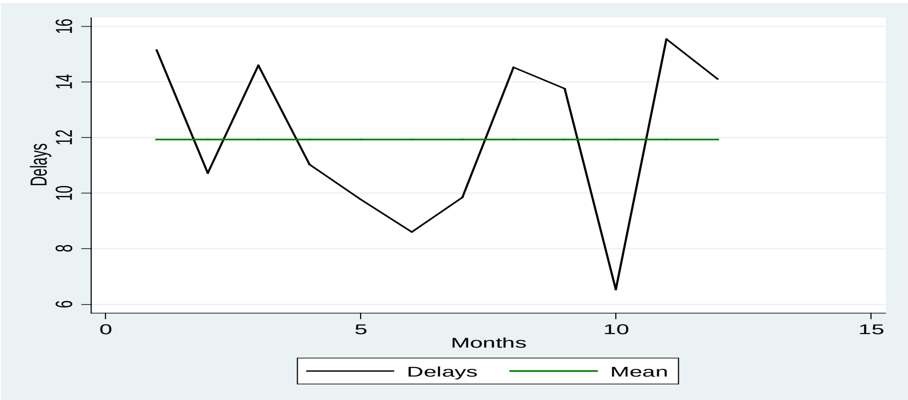

Figure 3 Duration of Delay in Minutes by Month

3.2 Analysis of variance for airline with day and month of the year

A Two-factor analysis of variance was conducted to evaluate the effect of day of the week on the delays of the different airlines (BAW, Eagle Air, KQA and SAA). The two independent variables were airline and day of the week. The dependent variable was the delays in minutes for the airlines.

Table 2: Two-way Analysis of Variance for the Delays, Airlines and Day of the week

Source	*SS*	*Df*	*MS*	*F*	*P*
Airline	328.98	3	109.66	162.62	0.0000

Day	12.72	6	2.12	3.14	0.0046
Airline X Day	14.67	16	0.92	1.36	0.1539
Residual	761.32	1129	0.67		
Total	1131.42	1154			

A significance level of 0.05 was used for the initial analyses. The results for the two-way ANOVA indicated a significant main effect for Airline F (3, 1129) = 162.62, p <0.001 and also a significant main effect for the Day of the week, F (6, 1129) = 3.14, p<0.05. Additionally, the results show a non-significant interaction between Airline and Day of the week, F (16, 1129) = 1.36, p>0.001 indicating that any differences in the average delays amongst the airlines were not dependent on the day of the week.

Approximately 29% (ω^2=.289) of the total variance of the delays was attributed to the airlines.

Since the interaction between airline and the day of the week were not significant, so the focus was turned on the main effect of the airline and day of week. Post hoc procedures of pairwise comparison were done on the Airlines and Day of week. A Sidak-test was preferred as it yielded only one non-significant difference between BAW and Eagle air (p>0.001) while the rest of the pair-wise comparison were significant. Further it was observed that there was no significant difference among the days of the week.

A Two-factor Analysis of Variance was conducted to evaluate the effect of month of year on the delay duration of the four different airlines (BAW, EA, KQA and SAA). The two independent variables considered here are airlines and months while the dependent variable was the airline delay in minutes.

Table 3: Two-way Analysis of Variance for the Delays, Airline and Month.

Source	*SS*	*Df*	*MS*	*F*	*P*
Airline	312.15	3	104.05	171.41	0.0000
Month	61.17	11	5.56	9.16	0.0000
Airline Month	62.05	33	1.88	3.10	0.0000
Residual	671.98	1107	0.6		
Total	1131.4	1154	0.98		

At a significance level of 0.05, the results for the two-way ANOVA indicated a significant main effects for Airline F (3, 1107) = 171.41, p <0.001 and for the Months F (11, 1107) = 35.56, p<0.001. Additionally, the results show a significant interaction between Airline and months, F (33, 1107) = 1.88, p<0.001 (See Table 3) indicating that any differences in the average delays amongst the airlines were dependent on the month of the year.

Approximately, 3.8% (ω^2=.038) of the total variance of the delays was attributed to the interaction between the airlines and the months while 4.8% (ω^2=0.048) of the total variance was attributed to the months of the year alone.

Since the interaction between airline and months of the year accounted for only 3.8%, we chose to ignore the two main effects and instead examined the airline simple effects, which is the difference among the airlines for each of the month. To control for type I error rate across the 12 months, we set the alpha level for each at 0.0042

($\alpha/12=.05/12$). There was significant difference among all the airlines across all the months under study implying that each airline is affected differently during a particular month.

We also examined the month simple effects, which were the differences among the 12 months for BAW, EA, KQA and SAA separately. To control for Type I error across the four simple effects, we set the alpha level for each at 0.0125 ($\alpha/4=.05/4$). There was a significant difference among the 12 months of the year for delay in the different airlines.

4. Discussions and Conclusions

The number of schedule flights had a positive effect on the departure delay [13, 14] implying that the number of schedules significantly affected the average delays by the airline as registered at Entebbe International Airport. Increase in schedules of flights is reflected in increased length of delay registered by the airline[14]. As per their findings in the United States airline industry [15] found out that increases in multimarket contact led to increases in delays.

This may have a number of policy implications; whereas the increased number of schedules at an airport is good as a source of income for airport management, in this case, it implies less efficiency registered. Thus, a policy framework needs to be developed to determine a maximum level of scheduling that would not compromise airport management income and operational departure efficiency.

The interaction between the days of the week and the airlines at the airport was not significant which implied that the effect of the day of the week on the delays did not depend on the type of airline. But it should further be noted that the days of the week has an effect on the average delays of the airlines [16].

The interaction between the month and the airlines at the airport was significant implying that the effect of the months on the delays was found to depend on the type of airline. Furthermore, it was noted from the posthoc tests of pair wise comparison that the different airlines were affected differently in the different months of the year. These differences in airline delay duration arising due to differences in the week or month could be associated to differences in meteorological parameters as explored by [5] in a paper parameterized framework for the analysis of probabilities of aircraft delay at an airport.

The study findings concur with [17-19] that the analysis framework can be used to improve the transportation planning and evaluation process. Each airline should be treated individually according to its unique characteristics in the process of improving performance and also that there is need to update the analysis in order to be able to represent the real system since a onetime analysis may not be sufficient. Given that airline delay is positively correlated with the number of scheduled flights, it is reflective of passenger choice of time of travel and calls for a policy framework to optimise schedules and airline delays during departure at the airport.

Acknowledgement

We extend our appreciation to the staff and management of the School of Statistics and Planning for the moral, material and financial support contributed towards this study. We also thank the management of the Civil Aviation Authority, Uganda and the Department of Meteorology for availing us with the data. Finally, we are highly indebted to the anonymous reviewers for the positive criticisms.

References

[1] Civil Aviation Authority, "Civil Aviation Authority, Uganda - Annual Statistics," Civil Aviation Authority, Uganda, Kampala 2012.

[2] Q. Gao, J. Yan, and J. Zhu, "Airlines' Optimization Decision of Slot Allocation in CDM," *Journal of Transportation Systems Engineering and Information Technology,* vol. 11, pp. 94-98, 2011.

[3] R. Wesonga, "Stochastic Optimization Models for Air Traffic Flow Management, Makerere University, Kampala.," in *School of Statistics and Planning*. vol. PhD (Statistics) Kampala: Makerere University, 2010, p. 200.

[4] G. Santos and M. Robin, "Determinants of Delays at European Airports," *Transportation Research Part B,* pp. 392-403, 2010.

[5] R. Wesonga, F. Nabugoomu, and P. Jehopio, "Parameterized framework for the analysis of probabilities of aircraft delay at an airport," *Journal of Air Transport Management,* vol. 23, pp. 1-4, 2012.

[6] M. Janic, *Airport Analysis, Planning, and Design: Demand, Capacity, and Congestion*: Nova Science Publishers, Incorporated - Transportation, 2009.

[7] P. P.-Y. Wu and K. Mengersen, "A review of models and model usage scenarios for an airport complex system," *Transportation Research Part A: Policy and Practice*, vol. 47, pp. 124-140, 2013.

[8] C. Gano and S. Banavar, "National Airspace System Delay Estimation Using Weather Weighted Traffic Counts," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*: American Institute of Aeronautics and Astronautics, 2005.

[9] S. Wang, Y. Zhang, Z. Zhang, and H. Yu, "Multi-objectives Optimization on Flights Landing Sequence at Busy Airport," *Journal of Transportation Systems Engineering and Information Technology,* vol. 12, pp. 135-142, 2012.

[10] M. Eric and C. Gano, "Analysis of Aircraft Arrival and Departure Delay Characteristics," in *AIAA's Aircraft Technology, Integration, & Operations (ATIO) 2002 Technical Forum*: American Institute of Aeronautics and Astronautics, 2002.

[11] N. Xu, L. Sherry, and K. Laskey, "Multifactor Model for Predicting Delays at U. S. Airports," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2052, pp. 62-71, 2008.

[12] E. M. Winter, R. G. Eston, and K. L. Lamb, "Statistical analyses in the physiology of exercise and kinanthropometry," *Journal of Sports Sciences*, vol. 19, pp. 761-775, 2013/06/30 2001.

[13] Y. Tu, M. O. Ball, and W. S. Jank, "Estimating Flight Departure Delay Distributions: A Statistical Approach With Long-Term Trend and Short-Term Pattern," *Journal of the American Statistical Association,* vol. 103, pp. 112-125, 2013/06/30 2008.

[14] Q. Cui, H.-b.Kuang, C.-y. Wu, and Y. Li, "Dynamic formation mechanism of airport competitiveness: The case of China," *Transportation Research Part A: Policy and Practice,* vol. 47, pp. 10-18, 2013.

[15] J. Prince and D. Simon, "Multimarket Contact and On-Time Performance in the U.S. Airline Industry," *Academy of Management Journal*, pp. 336-354, 2010.

[16] L. A. Zadeh, "Generalized theory of uncertainty (GTU)” principal concepts and ideas," *Computational Statistics & Data Analysis*, vol. 51, pp. 15-46, 2006.

[17] C. Hendrickson and G. Kocur, "Schedule Delay and Departure Time Decisions in a Deterministic Model," *Transportation Science,* vol. 15, pp. 62-77, February 1, 1981 1981.

[18] A. Sasic and K. N. Habib, "Modelling departure time choices by a Heteroskedastic Generalized Logit (Het-GenL) model: An investigation on home-based commuting trips in the Greater Toronto and Hamilton Area (GTHA)," *Transportation Research Part A: Policy and Practice,* vol. 50, pp. 15-32, 2013.

[19] C. Speier, J. M. Whipple, D. J. Closs, and M. D. Voss, "Global supply chain design considerations: Mitigating product safety and security risks," *Journal of Operations Management,* vol. 29, pp. 721-736, 2011.