Journal of Air Transport Management 23 (2012) 1–4

Contents lists available at SciVerse ScienceDirect

Journal of Air Transport Management

journal homepage: www.elsevier.com/locate/jairtraman

Parameterized framework for the analysis of probabilities of aircraft delay at an airport

Ronald Wesonga [a,*], Fabian Nabugoomu [b], Peter Jehopio [a]

[a] School of Statistics and Applied Economics, Makerere University, PO. Box 7062, Kampala. Uganda
[b] Uganda Christian University, Mukono, Uganda

Keywords:
Aircraft delay
Probability of delay
Threshold levels of airport delays

A B S T R A C T

The study analyses ground delays and air holding at Entebbe International Airport over five years. Daily probabilities for aircraft departure and arrival delays at are generated for each. The mean probabilities of delay for ground delays and air holding at 50% delay threshold levels are 0.94 and 0.82 that fall to 0.49 and 0.36 when 60% delay threshold levels are used. Simulations are performance for delay threshold levels to monitor for the trends of the daily probabilities for the study period. The general conclusion is that a parameter-based framework is best suited to determine the probability of aircraft delay at an airport.

© 2012 Elsevier Ltd. All rights reserved.

1. Background

The time arrival patterns of aircraft largely, but not exclusively, determine airport delay for departing aircraft. Mueller and Chatterji (2002) established that determining the probability of aircraft delay at an airport is, however, challenging. Tu et al. (2008) and others have attempted to model flight departure delay probability distributions, but without assessing the totality of the determining aviation and meteorological parameters. Aircraft delay is measured as the difference between the estimated and actual time of an aircraft's departure or arrival, delays. When an aircraft departs or arrives earlier than expected this is generally treated as an on time operation. Here we derive a parameter-based probability model of aircraft delay for any given day at an airport.

2. Methodology

This study involves two random aircraft operations that occur at an airport, departure and arrival. Unlike (Markovic et al., 2008) who analysed aircraft punctuality focusing on weather parameters, aircraft delay forms the epicentre of our analysis. Aircraft delays are divided into ground (GDP) and air holding (AHP) delays. Given that on any given day, the number of aircraft that are delayed either departing or arriving can be established, days can classified as either delay or on-time day based on whether delays reached a threshold level.

A logistic model is applied where the dependent dummy variable (y_i) is zero if during a given day aircraft' operations are classified as being on time and one if they are delayed. The estimation of the probability of a delay is done using the cumulative logistic distribution, where the β's are coefficients of the x_{ij} categories to which the sample unit belongs.

$$\ln(\pi(X_i)) = \sum_{j=1}^{p} \beta_j x_{ij} \tag{1}$$

where $\ln(\pi(X_i))$ represents the logarithm the conditional probability that a certain day is classified as a delay day given all explanatory variables, and its determinants are subsequently tested for significance of the underlying relationship.

$$\text{Odds} = \frac{\pi(X)}{1-\pi(X)} = \exp^{\sum_{j=1}^{p} \beta_j x_{ij}} \tag{2}$$

This implies that the odds are exponential function of X_i that provides a basic interpretation of the magnitude of the coefficients. When β_j is positive, it implies increasing rate while when β_j is negative, this implies decreasing rate and the rate of climb or descent increases as the magnitude of β_j increases. Conversely, the magnitude of β_j signifies the increasing or decreasing effect of a given determinant on the daily proportion of delay. If $\beta_j = 0$, it would mean that the daily proportion of aircraft delay is independent of X_i.

$$\pi(X_i) \frac{\exp^{\sum_{j=1}^{p} \beta_j x_{ij}}}{1 + \exp^{\sum_{j=1}^{p} \beta_j x_{ij}}} \tag{3}$$

where $\pi(X_i)$ is the probability that on a given day the proportion of aircraft that are delayed in departure or arrival given the influence of the meteorological and aviation parameters as shown in Table 1.

* Corresponding author. Tel.: +256 772454680.
E-mail address: wesonga@wesonga.com (R. Wesonga).

0969-6997/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jairtraman.2012.02.001

Table 1
Aircraft delay based parameters.

Field name	Type	Upper limit	Lower limit	Continuous/Discrete	Description
Scheduled	Integer	Dec. 2008	Jan. 2004	Discrete	Daily scheduled flights
Non-scheduled	Integer	Dec. 2008	Jan. 2004	Discrete	Daily non-scheduled flights
Domestic	Integer	Dec. 2008	Jan. 2004	Discrete	Daily domestic flights
International	Integer	Dec. 2008	Jan. 2004	Discrete	Daily international flights
POBin	Integer	Dec. 2008	Jan. 2004	Discrete	Daily persons on board on the incoming aircraft
POBout	Integer	Dec. 2008	Jan. 2004	Discrete	Persons on board on the outgoing aircraft
GDP	Integer	Dec. 2008	Jan. 2004	Discrete	Proportion of aircraft that have delayed to depart on a given day
AHP	Integer	Dec. 2008	Jan. 2004	Discrete	Proportion of aircraft that have delayed to arrive on a given day
Visibility	Float	Dec. 2008	Jan. 2004	Continuous	Average daily visibility
Windrecn	Float	Dec. 2008	Jan. 2004	Continuous	Average daily visibility
Windspeed	Float	Dec. 2008	Jan. 2004	Continuous	Average wind speed
QNH	Float	Dec. 2008	Jan. 2004	Continuous	Queen's nautical height

The idea of an aircraft delay threshold level is introduced to reflect the daily proportions of aircraft delay that contribute most to overall delays as measured by statistically significant explanatory factors. An aircraft delay threshold level is defined as the proportion beyond which a given day is classified as a delay day or on-time day. Data simulations using threshold levels for GDP and AHP at an airport are used.

3. Classifying airport days

All calendar days are assumed to be airport days where normal operations occur on all days. What varies is the number of aircraft that depart and arrive within a day. As a result the proportions of aircrafts that delay to depart or arrive also vary from one day to the other. The process of classifying an airport day is totally a random variable based only on the timeliness operation of aircraft departures or arrivals at an airport. Let's assume different number of aircraft depart on different days. Let the N represent the number of aircraft and M number of operational days at an airport and when an aircraft departs later than the scheduled time, then it is delayed or else it departs on time. Thus we have the data matrix;

$$\begin{array}{lllll} \text{Day}_1 & A_{11}, & A_{12}, & \ldots, & A_{1N} \\ \text{Day}_m & A_{m1}, & A_{m2}, & \ldots, & A_{mN} \\ \text{Day}_M & A_{M1}, & A_{M2}, & \ldots, & A_{MN} \end{array}$$

where, M, the number of days, can be fixed based on the time period, but N, the number of aircraft that depart and arrive in a given day may differ.

4. Logistic model for air traffic delay

A logistic model using a two-category dummy variable reflecting the proportion of aircraft delaying their operations and the proportion operating on time is used generating probabilities for an aircraft operating on time based on daily delay patterns (Table 2). The table shows parameters based on the occurrence of a 50% threshold of aircraft experiencing delay at departure and arrival at Entebbe International Airport. We see that the number of freighter movements and non-commercial flights per day significantly influence both arrival and departures delays. Other factors that significantly affect the proportion of departure delays on a certain day given a 50% threshold are the arrival delay dummy, and the number of arrival delays, operations, scheduled flights and chartered flights per day. The odds ratios (OR) show that the parameters; proportion of arrival delay and number of operations unit increases would result into a reduction of in the proportion of departure delay by 89% and 68%.

To explain the proportion of arrival delays, the variables found to be significant, but with negative coefficients were the departure delay dummy and number of freighters. The rate of their effect shows that a unit increase in departure delay would result into a 57% reduction of the proportion of arrival delays compared to 87% reduction from a unit increase in the number of freighters. Since the logistic model is of sigmoid form, this implies that the rate of change in the odds, $\pi(X_i)$ per unit change in the explanatory variables x_i varies according to the relationship $\partial\pi(X_i)/\partial(x) = \beta\pi(X_i)[1-\pi(X_i)]$. Thus for the odds of the proportion of delay $\pi(X_i) = 0.5$, and taking the coefficient of the number of scheduled flights, $\beta = 0.46$ the slope is$\partial\pi(X_i)/\partial(x_i) = 0.46{*}(1/2){*}(1/2) = 0.115$ representing a change in the odds of departure delay, $\pi(X_i)$ per unit change in the number of scheduled flights. Put another way, for every 100 scheduled flights departures at Entebbe International Airport, eleven flights experience departure delay given a threshold level of 50%.

Post logistic estimation analysis was performed to estimate the probability of the daily proportions of aircraft at departure and arrival delay by computing mean values of the estimated daily probabilities. This analysis generated daily probabilities for the 1827 days within the period. The overall average probability for departure and arrival delay over the period is estimated (Table 3).

Generally, holding other parameters constant at 50% threshold level, the probability of aircraft departure delay was established to be relatively higher than the probability for aircraft arrival delay. Based on the collaborative nature of airports, one can conclude that

Table 2
Logistic model dynamics for aircraft departure and arrival delay.

DV: dummy daily proportion of departure delay	Odds ratios	DV: dummy daily proportion of arrival delay	Odds ratios
Intercept	3.00*	Intercept	2.59**
Arrival delay dummy	2.05*	Departure delay dummy	0.57*
Proportion of arrival delay	0.89**	Number of freighters	0.87**
Number of operations	0.68**	Number of other non-commercial flights	1.03**
Number of scheduled flights	1.58**	Number of persons on board in	1.01**
Number of chartered flights	1.38*		
Number of freighters	1.80**		
Number of other non-commercial flights	1.77**		

Note: ** significant at 0.01 level; * significant at 0.05 level.

Table 3
Probability of aircraft departure and arrival delay.

Category	1st quartile	3rd quartile	Mean probability
Departure delay given a 50% delay threshold	0.92	0.99	0.94
Arrival delay given a 50% delay threshold	0.77	0.89	0.82

Table 4
Variation of predicted delay probability with the threshold level.

Delay threshold level (%)	Probability using logistic model							
	Departure delay				Arrival delay			
	Variables in the model	1st quartile	3rd quartile	Mean	Variables in the model	1st quartile	3rd quartile	Mean
50	8	0.92	0.99	0.94	4	0.77	0.89	0.82
60	9	0.17	0.83	0.49	10	0.12	0.55	0.36
70	7	0.02	0.50	0.26	3	0.01	0.32	0.18
80	2	0.00	0.08	0.05	3	0.00	0.05	0.04

the lower arrival delay is mainly due to factors that are exogenously determined outside the Airport, Entebbe International Airport.

Furthermore, post logistic estimation analysis for threshold levels of {50, 60, 70, 80} were performed to estimate the probability of departure and arrival delay (Wesonga, 2010). Use of different threshold levels generates dependent variables with varying counts of statistically significant parameters (Zimmermann, 2008). The lower threshold values {40, 30, 20, 10} were not used because they generate fewer counts for the second category that would not work for the logistic modelling. The results are shown in Table 4.

The results show that predicted delays for aircraft at departure and arrival reduce as the threshold level increase. The mean predicted probability for aircraft departures and arrivals at the airport that used more predictors were 0.49 with 9 predictors and 0.36 with 10 predictors respectively. However, in both of these cases, there is a characteristic large deviation between the estimates for the 1st and 3rd quartiles, but larger for the departure than arrival estimated probabilities.

5. Probabilities from the logistic models

Post logistic modelling analyses for the behaviour of the probabilities covering 1827 records was done. It is evident from Fig. 1 that the lower the threshold proportion of delay; the greater are the estimated probabilities that the airport will experience a departure delay. Further, as the threshold is increased allowing less departure delay the predicted probabilities over time smoothen at the 80% threshold. Generally, it indicated that predicted aircraft departure delay probabilities exhibited a negative trend between 2004 and 2008.

Fig. 2 illustrates that the lower the delay threshold, the greater are the probabilities that the airport will experience arrival delay. As the value of the threshold increases thereby allowing smaller arrival delays so the probabilities over time breaks into two trends from 2007 at 60% and 70% threshold levels, but this almost smoothens at the 80% level indicating that the aircraft arrival delay probabilities exhibited a slightly positive trend over the period 2004 to 2008.

Probabilities of departure and arrival delay were computed annually using the threshold with more predictors of 60% for both GDP and AHP. Table 5 shows how the probabilities of delay varied over years. It was noted that these probabilities are conditional on parameters influencing delay at the Airport, Entebbe International Airport.

The collaborative nature of air traffic flow management at different airports means that an aircraft's arrival performance may be due to parameters outside the arrival airport. The complexity of this collaboration is premised on the fact that any arriving aircraft must have departed from some other airport. Therefore, the timeliness of the arriving aircraft is affected not only by parameters at the arrival airport, but also others that are exogenously determined especially at the departing airports. Similarly, departing aircraft are

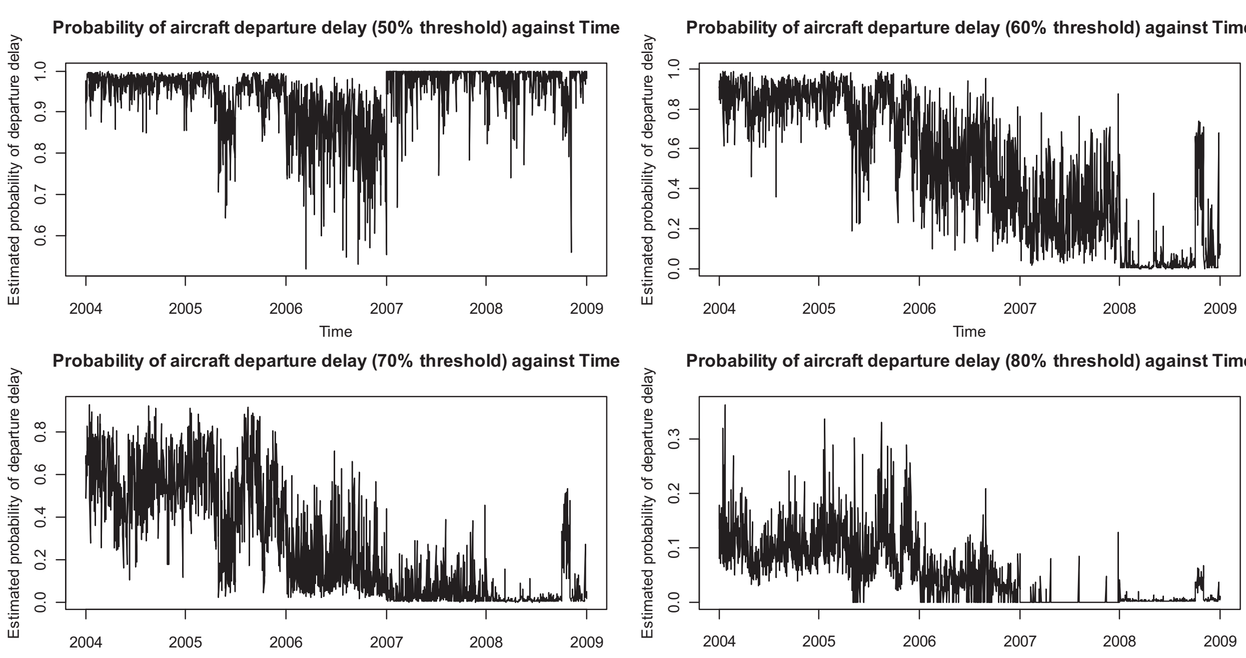

Fig. 1. Variation of predicted departure delay probability with time (days).

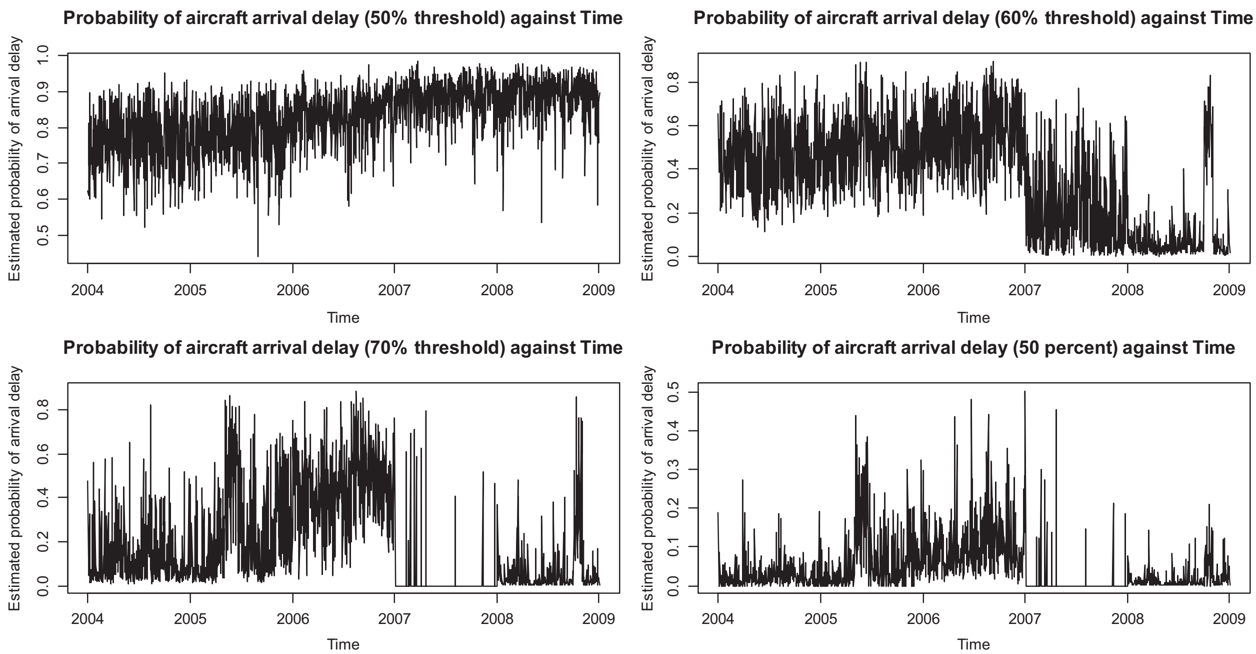

Fig. 2. Variation of predicted arrival delay probability with time (days).

Table 5
Variation of probability of departure and arrival delay from 2004 to 2008.

Year	Probability of departure delay	Probability of arrival delay
2004	0.9454	0.3443
2005	0.8986	0.4055
2006	0.3534	0.9589
2007	0.2164	0.0931
2008	0.0795	0.0164

Table 6
Aircraft delays at 60% threshold level.

Year	Ground delay programme (GDP)			Air holding programme (AHP)		
	Minimum	Maximum	Mean	Minimum	Maximum	Mean
2004	0	52	28	0	43	10
2005	0	47	24	0	40	10
2006	0	39	8	0	56	29
2007	0	39	5	0	40	2
2008	0	48	2	0	32	1

primarily determined by parameters within the airport and also those at other airports.

Descriptive statistics for the delayed aircrafts under the 60% threshold level are shown in Table 6. The analysis indicated that there was a decreasing average number of aircraft delay between 2004 and 2008 for both departure delay (GDP) and arrival delay (AHP). The analysis was done at the 60% threshold because it is at this level that the most number of explanatory parameters was established as shown in Table 2.

6. Conclusions

This study uses a multiple parametric approach to determine the probability of aircraft delay. It uses a robust approach to include the apparently significant meteorological and aviation parameters while computing the exact probabilities of delay. Significant parameters were assessed at delay threshold levels for Entebbe International Airport. These levels of significance are expected to vary between airports depending on measurements of aviation and meteorological parameters. In terms of an appropriate aircraft delay threshold level for the case study, 60% was found to have the largest number of significant parameters affecting both departure and arrival delay estimates. The probabilities of aircraft delay in either case were 49% and 36% suggesting that a departing aircraft is more likely to delay compared to one arriving. Furthermore, aggregated annual probabilities of delay over 2004 through 2008 show a decline in the probability of delay.

References

Markovic, D., Hauf, T., et al., 2008. A statistical study of the weather impact on punctuality at Frankfurt airport. Royal Meteorological Society 15, 293–303.

Mueller, E.R., Chatterji, G.B., 2002. Analysis of Aircraft Arrival and Departure Delay Characteristics AIAA's Aircraft Technology, Integration, and Operations (ATIO) 2002 Technical. NASA Ames Research Center, Moffett Field, Los Angeles.

Tu, Y., Ball, M., et al., 2008. Estimating flight departure delay distributions—a statistical approach with long-term trend and short-term pattern. Journal of the American Statistical Association 103, 112–125.

Wesonga, R., 2010. Stochastic optimisation models for air traffic flow management. Planning and Applied Statistics. Uganda, PhD Thesis Makerere University.

Zimmermann, A., 2008. Stochastic Discrete Event Systems: Modeling, Evaluation, Applications. Springer, New York.