

On the Modeling of Airport Arrival and Departure Delay Distributions

Qian Wu

School of Information Engineering, MinZu University of China, Beijing, China, P.R

wuqian@muc.edu.cn

Abstract

Airport arrival and delay distribution characteristics are analyzed and then modeled using probability density functions against a collection of departure and arrival delay data over an 8-month period of Beijing Capital International Airport (BCIA). We present an optimized Generalized Extreme Value (GEV) model in the paper and demonstrate through application of BCIA's history flight delay data better Goodness of Fit. The quantitative modeling of delay characterizations in terms of stochastic distributions would play an important role in improving demand predictions in air traffic flow management systems.

Keywords: Airport Delay Characteristics, Air Traffic Flow Management, Generalized Extreme Value, Probability Density Function.

1 Introduction

To tackle with the soaring demands of air transportations in China and around the world, academic research and industrial practices have been promoting toward a collaborative and integrated air traffic flow management infrastructure, to improve the efficiency and performance of all stakeholders of national air space, Nowadays the air traffic flow management aiming at coordinating airspace resources and streamlining air traffic flow, to ensure traffic demands in line with capacities, has become the most widely adopted approach for reducing airspace congestions, since its initial concepts and practices emerging from mid 1990's in both America and Europe[1]. Among other criteria, the accuracy of traffic demand

prediction almost becomes the precondition for nearly all effective ATFM operations. Intuitively, air traffic prediction relies upon predictions of number of aircrafts in a given region based on aircraft trajectories. However, uncertainties over the air traffic operations such as weather conditions often cause randomness in departure, arrival or airborne delays. In many practices, for example, departure time uncertainty could be a major cause of demand forecast discrepancy [2]. To facilitate ATFM operations, the effective and accurate modeling of airport delay characteristics has become one of key factors to improve the accuracy of traffic demand predictions.

Many stochastic modeling techniques have been published over the past decade. Related work [3][4] revealed a modeling technique using Normal and Poisson distributions. The raw airport delay data over a 21-day period collected among ten major US airports were analyzed to derive the mean and standard deviations, then to construct a set of delay-time probability density functions. Whereas in another paper [5], an aggregate stochastic model was developed using a stochastic description of an air traffic system based on Poisson processes, to predict dynamic aircraft counts in regions of the airspace selected. In related work [6], however, based on air traffic data over two regions in U.S. airspace, Eulerian models and aggregated traffic models were used to represent air traffic environment instead.

We in this paper nevertheless emphasize the modeling of stochastic processes on a single airport arrival and departure delays using an optimal Generalized Extreme Value (GEV) model. Similarly we collected historical traffic data of BCIA over an eight months period from January to August 2012 as basis to characterize the probability density distributions for departure and arrival delays. Compare with other models, the resulted departure and arrival delay probability density functions of BCIA are derived based on proposed optimal GEV model, with better Goodness of Fit shown through application of real data.

The paper will present in the following sections. In section 2 the raw data preparation for our analysis is illustrated. Then in section 3, the model formulation

using an optimal Generalized Extreme Value (GEV) is present, followed by the model verification through application of real data. We conclude the paper and anticipate the future work in section 4.

2 Airport Delay Data Preparations

All historical data in this paper were collected from actual flight data of BCIA during the period from January to August 2012. Flight data fields were extracted to include just flight date, flight number, source and destination airport, scheduled time of departure, actual time of departure, scheduled time of arrival and actual time of arrival. The raw data were firstly filtered to include BCIA as either source or destination airport only. To better understand delay characteristics, it is useful to pre-calculate the average daily departure/arrival delay distribution for BCIA during the period from January to August 2012. Results are listed in Table 1. Table 1 shows average daily departure delay time is greater than average daily arrival delay. The time difference is about 30 minutes..

Table 1: Daily Average Delay Statistics of BCIA over an 8-month period

Month	Ave. Daily Delay(min)			Std Dev,of Delays (min)		Ave. Daily Flights	
	Dep	Arr	Diff	Dep	Arr	Dep	Arr
2012-01	41.73	12.23	29.5	61.35	72.92	718	667
2012-02	36.04	7.79	28.25	49.21	61.6	787	738
2012-03	42.89	13.73	29.16	62.9	68.35	734	687
2012-04	51.01	19.76	31.25	70.7	79.4	736	670
2012-05	41.05	11.93	29.12	54.25	62.7	741	688
2012-06	56.05	22.44	33.61	69.29	76.86	742	671
2012-07	63.85	27.16	36.69	78.65	80.75	755	687
2012-08	52.14	21.97	30.17	60.41	73.58	789	764
Total	48.10	17.13	30.97	63.35	72.02	750	697

Further illustrations on the monthly statistics of average daily departure and arrival delay over the period from Jan 2012 to Aug 2012 , as depicted in Figure 1.



Fig. 1 Monthly Statistics of Average Daily Delay (min) 2012/01-08

3 Model Formulations

A. Gegneralized Extreme Value (GEV)

The generalized extreme value (GEV) is mainly used to describe the tail characteristics of the distribution. Let $X_1 \dots X_n$ be independent random variables following the GEV distribution, which is depicted as follow:

$$F(x) = p(x \leq X) = e^{-\left(1+k\left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{k}}\right)}, k \neq 0;$$

Where,

μ : the position coefficient;

σ : the scale coefficient, and

k : the envelop coefficient.

The probability density function of the GEV distribution is:

$$y = f(x | k, \mu, \sigma) = \left(\frac{1}{\sigma}\right) e^{-\left(1+k\left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{k}}\right)} \left(1+k\left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{k}}\right)^{-1-\frac{1}{k}}, \text{ where -}$$

$$1 + k\left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{k}} > 0. \text{ and}$$

When $k > 0$, it is extremal type I; $k < 0$ it is extremal Type II, and when $k = 0$, it is extremal Type III.

B. GEV Distribution Model

Importing delay data of Aug. 2012 into the Matlab, we adopted the “maximum likelihood estimation” (MLE) to estimate three coefficients of the GEV model. Figure 2(a) shows the “GEV model” distribution fitting curve (blue), in comparison with the raw density distribution curve. It shows that there is a little difference on the peak. To minimize the error of the raw distribution data, the envelop coefficient k and scale coefficient σ value can be adjusted based on the MLE estimation to create an Optimized GEV distribution model as shown in figure 2(b).

Obviously, the “Optimized GEV Model” better fits the raw density distribution curve (black) as shown in figure 2 (b). Similar results were obtained for modeling arrival delay density function.

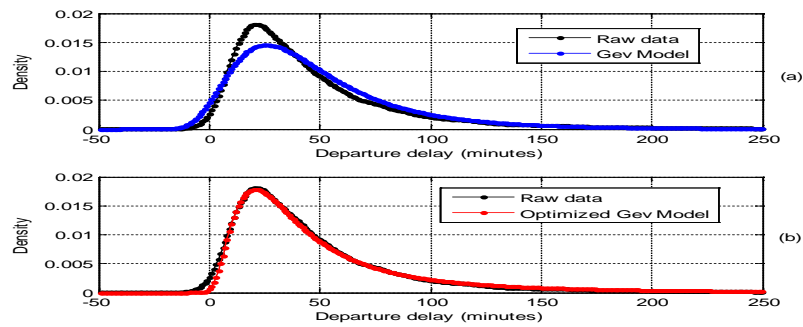


Fig.2 GEV Modeling of departure delay of August 2012

It is shown in Figure3(a) that the GEV Model has distinct differences with raw distribution, whereas the Optimized GEV Model in figure3(b) fits the raw density

distribution very well.

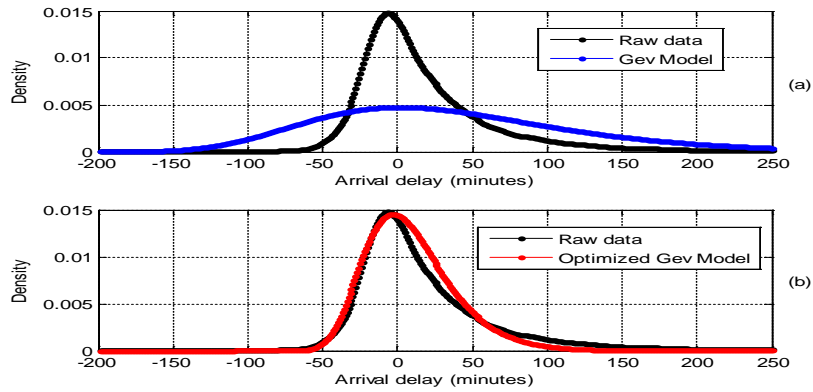


Fig.3 (a) The GEV model and (b) Optimized GEV model of arrival delays in August 2012

Table 2 shows the value of coefficients used for “GEV Model” and “Optimized GEV Model”,

RMSE (Minimum Mean Square Error) for estimating the fitting error is also listed in table2. “Optimized GEV Model” shows an improvement of RMSE about 50% over “GEV Model” for departure delays. The envelop coeff. $K > 0$ indicates the departure delay follows the extremal Type III distribution. Similarly, The Optimized GEV Model improves up to 77% over the GEV Model in fitting error estimations; and the fact that $k < 0$, means arrival delays follows the extremal type II distribution.

Fitting Model			RMSE	
	Departur e Delays	Arrival Delays	Departure Delays	Arrival Delays
GEV Model	$k =$ 0.1767	$k =$ -0.054 4	5.414e-04	1.3e-03
	$\sigma =$ 25.68	$\sigma =$ 77.28		
	$\mu =$ 29.81	$\mu =$		

		-2.7		
Optimized GEV Model	k = 0.5	k = -0.054 4	2.1674e-0 4	3.2408e-04
	$\sigma = 23$	$\sigma =$ 25.28		
	$\mu = 29$	$\mu =$ -4.7		

Table 2: Coefficients Values and RMSE for “GEV Model” and “Optimized GEV Model” for Departure/Arrival Delays

4 Model Verifications

To verify the optimized GEV model, we collected raw departure delay data of August 2012 as samples. Comparing the difference between raw departure data statistical results and GEV model result, the minor difference indicates a better model.

The probability distribution function for raw departure delay data is compared with the output of optimized GEV model. The result shown in Figure 4 demonstrates that envelopes of these two curves are very fitting well.

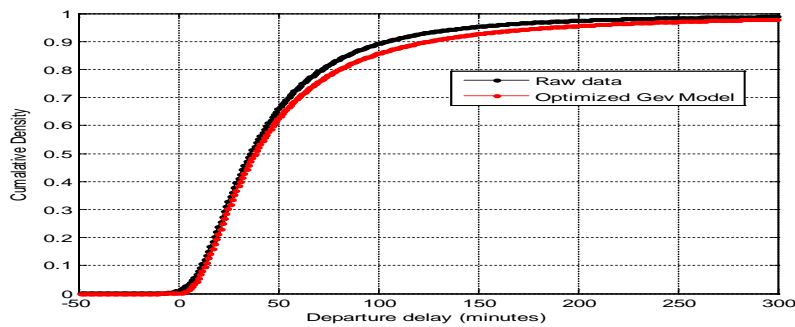


Fig.4 Probability Distribution Function for raw departure delay and optimized GEV model output

5 Conclusions

The paper adopted the GEV model to characterize the departure and arrival delay of BCIA based upon the data collected from BCIA operations from Jan. to Aug. 2012. The Probability Density Functions of departure and arrival delay of BCIA are constructed using an optimal GEV model. The Goodness of Fit has been shown through application of real data. We will seek to refine the model further by applying raw data from other airports, and look forward to applying the model in ATFM decision making processes, in the hope that the delay characterizations in terms of stochastic distributions drawn from our work could improve demand predictions in air traffic flow management systems.

References

- [1] Federal Aviation Administration, "Monthly Summary Air Traffic Activity and Delays in the National Airspace System March 2001," www.faa.gov/apa/mar01del.htm, March 2001
- [2] Menon, P. K., Sweriduk, G. D., and Bilimoria, K. D., "A New Approach for Modeling, Analysis and Control of Air Traffic Flow," *AIAA Guidance, Navigation and Control Conference*, August 5 – 8, 2002, Monterey, CA
- [3] Robinson J E III, and Isaacson D R. "A Concurrent Sequencing and Deconfliction Algorithm for Terminal Area Air Traffic Control", AIAA Paper 00-4473, AIAA Guidance, Navigation, and Control Conference, Denver, Colorado, 2000
- [4] P. K. Menon, G. D. Sweriduk, and B. Sridhar. Optimal strategies for free flight air traffic conflict resolution. *Journal of Guidance, Control, and Dynamics*, 22(2):202–211, 1999
- [5] Mueller, E., and Chatterji, G., "Analysis of Aircraft Arrival and Departure Delay Characteristic," AIAA Paper 2002-5866, Oct. 2002.
- [6] Roy, S, Sridhar, and Verghese, G.C., "An Aggregate Dynamic Stochastic Model for an Air Traffic System," in Proc. of the 5th Eurocontrol/Federal

Aviation Agency Air Traffic Management Research and Development
Seminar, Budapest, Hungary, June, 2003.