Journal of Air Transport Management 52 (2016) 35–41

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

Analysis of short-term forecasting for flight arrival time

Myung Suk Kim

Sogang Business School, Sogang University, #1 Shinsu-dong, Mapo-gu, Seoul 121-742, Republic of Korea

A R T I C L E I N F O

Article history:
Received 18 June 2015
Received in revised form
9 November 2015
Accepted 17 December 2015
Available online 2 January 2016

Keywords:
Short-term forecasting
Air transport
Skew t distribution
Spline smoothing

A B S T R A C T

We suggest various methodologies to provide short-term forecasting of flight arrival times. Flights arriving at Denver International Airport from various U.S. cities during 2010 are used for the model estimation, and the forecasting is applied to 2011 flights. Forecasting proceeds from the time at which a flight departs from an airport. Prediction models using the spline smoothing-based nonparametric additive techniques are applied and compared with benchmarks. We also provide a method for computing the probability of flight arrival time by fitting the skew t distribution to the models' residuals. Our empirical results indicate that a nonparametric additive model dominantly outperforms the other models considered. In terms of effect of predictor variables, departure delay time, scheduled airborne time, airlines, and weather conditions significantly improve forecasting accuracy, along with seasonal variables. In particular, departure delay time is the most important factor for substantially improving prediction performance.

1. Introduction

Modeling and forecasting flight departure and arrival time have been significant issues for both industry and academia. Recently, U.S. commercial airlines spent approximately $20 billion due to flight delays (Schumer and Maloney, 2008). Some rigorous research for costs to airlines from delays was carried out (Ball et al., 2010), and various attempts have been made to reduce costs. For example, Kaggle hosted the GE Flight Quest, at which more than 155 teams competed, using their algorithms to improve the prediction of delays (see https://www.gequest.com/c/flight).

Academic researchers also suggested various approaches to flight time modeling and forecasting. In the beginning, statistical linear or nonlinear type models were proposed and their prediction performances were evaluated. Xu et al. (2008) predicted positive and negative delays using multivariate adaptive regression spline models, which were useful in detecting a nonlinear relationship between the explanatory and response variables. Srivastava (2011) applied a linear regression model to predict taxi-out or taxi-out delay using various explanatory variables such as runway distance, queue position, arrival and departure rates, and weather. Rebollo and Balakrishnam (2012) proposed random forest algorithm-based models for air traffic delay prediction.

Contemporarily, in addition to developing flight time prediction models, some researchers were interested in computing the probability of flight delays. For example, Mueller and Chatterji (2002) proposed computing the probability of departure and arrival delays using Poisson and normal distributions, respectively. Tu et al. (2008) suggested a prediction model using daily propagation effects and seasonal trends for forecasting flight departure delays using a spline smoothing-based nonparametric additive approach. Meanwhile, parametric mixture distribution was applied to model residual errors, which were utilized to compute the probability of a delay. Deshpande and Arıkan (2012) suggested forecasting the truncated total travel time using various predictor variables such as routes, carrier, origin and destination airports, congestion, and aircraft-specific variables. They assumed that travel time follows log-normal and log-Laplace distributions, and calculated on-time arrival probability under these distributions.

In the current study, prediction models with nonparametric additive techniques are applied to the short-term forecasting of flight arrival time, which is extended from the approaches by Tu et al. (2008). Flights arriving data at Denver International Airport in 2010 and 2011 from various U.S. cities are used to estimate and forecast the model. Various predictor variables related to airborne state, and departing and arriving airports are implemented. Linear regression and median regression are also employed as benchmark models. The predictability performance between models is evaluated using point-wise prediction criteria. Lastly, the residual errors are modeled using the skew t distribution, which is utilized to compute the probability of the arrival delay.

E-mail address: myungkim@sogang.ac.kr.

http://dx.doi.org/10.1016/j.jairtraman.2015.12.002
0969-6997/

Our research findings make two important contributions to the literature. First, they show that departure delay time is very important in providing substantial improvements in the prediction accuracy of flight arrival time. In our analysis frame, arrival delay forecasting proceeds from the point at which a flight takes off from a departing airport. Therefore, departure delay information is available in the prediction. Although the forecasting horizon is very short, the newly collected delay time information can be utilized to update the prediction outputs depending on the availability of real-time business intelligence. Eventually, improved prediction may lead to more efficient air traffic control at the arriving airport. Second, we found that the skew t distribution is statistically appropriate to model the fitted residuals. Existing distributions were shown to be invalid through goodness-of-fit test, or no rigorous statistical goodness-of-fit test was implemented. Therefore, accuracy of computing the probability of arrival delay will improve.

The remainder of this article is organized as follows. In Section 2, data and variables are explained. The forecasting models are suggested in Section 3, wherein the nonparametric additive models and their benchmark models are described. Section 4 reports the forecasting results of each model. The arrival time probability using the skew t distribution is computed in Section 5. Concluding remarks are given in Section 6.

2. Data and variables

U.S. air carriers whose customers represent at least one percent of total domestic scheduled-service passengers are supposed to report their airline on-time data every month to the U.S. Department of Transportation (DOT) and the Bureau of Transportation Statistics (BTS). The data cover all nonstop scheduled-service flights between areas within the United States. This study employs DOT data on domestic air flights arriving at Denver International Airport from more than 100 cities in the United States. Because our approach focuses on short-term forecasting, this analysis excludes international flight arrivals. From a list of the top ten airlines in the United States based on 2010 passenger numbers, the top seven air carriers were selected, comprising Delta, United, Southwest, American, US Airways, Trans World Airlines, and SkyWest. Complete historical flight data of seven air carriers for all of 2010 were implemented in the model estimation, whereas forecasting performance was evaluated using data for all of 2011. Cancelled and diverted cases were deleted from the data set. A total of 187,514 and 171,971 air flights arriving at Denver International Airport were applied for 2010 and 2011, respectively.

2.1. Response variable

According to the DOT, a flight is classified as delayed if it arrives at the departing or arriving gate 15 min later than its scheduled time. However, no globally standardized method exists to measure flight delay. Flight delay was modeled using numerous methods. For example, Idris et al. (2002) and Srivastava (2011) considered the taxi-out delay, which is the duration between pushback and takeoff. Mueller and Chatterji (2002) and Tu et al. (2008) applied the pushback delay that is measured using the difference between the scheduled departure time and the actual departure time. Xu et al. (2008) employed positive and negative delays using the difference between scheduled arrival time and actual arrival time in some phases. Deshpande and Arıkan (2012) applied truncated total travel time delay that is measured using the discrepancy between departure time and arrival time.

In our model, the arrival delay is considered. Arrival delay is actual arrival time minus scheduled arrival time. Thus, positive, negative, or zero values are available. A negative computed arrival delay implies an early arrival. The descriptive statistics for arrival delay for 2010 are summarized in Table 1, which shows that their distributions are heavily skewed to the right, similar to the results of previous studies (Mueller and Chatterji, 2002; Tu et al., 2008).

Some authors transformed the target response variable to reduce the heteroskedasticity problem. For example, Deshpande and Arıkan (2012) took a log-transformation to the truncated total travel time. We attempted to employ the log-transformation, but it was not applicable to the negative values of the response variable. Therefore, before the log-transformation, we transformed all the possible negative values into positive ones by adding 150. Note that the minimum value of arrival delay is -81. Thus, 150 is big enough to switch the negative value cases into positive ones. That is, the original arrival delay is notated as N_i, whereas the forecasting will proceed using $y_i = \ln(N_i + 150)$.

2.2. Predictor variables

According to the BTS regarding all air carriers arriving at and departing from Denver International Airport in 2010, the percentage of flight delays reached approximately 16.4% of total flights. The detailed airline on-time statistics and delays are summarized on the BTS website (see http://www.transtats.bts.gov/ot_delay/ot_delaycause1.asp), which reports statistics for flight delay causes: air carrier delay (3.48%), aircraft arriving late (6.61%), security delay (0.04%), national aviation system delay (4.38%), extreme weather (0.42%), and cancelled and diverted (1.43%).

In this study, the causes of flight delay were reorganized into three groups based on the sources of arrival delays or early arrivals. These three sources are departing airport, airborne state, and arriving airport, which are summarized in Table 2.

Arrival delays are often caused by departure delays at the departing airport. Many factors influence departure time. For example, the congestion period and a relatively busy airport are seasonal and regional factors that affect airport capacity. Mayer and Sinai (2003) pointed out that air delays are influenced by airline hub size and airport concentration. Deshpande and Arıkan (2012) also considered how a hub airport affects flight delays. Rebollo and Balakrishnam (2012) claimed that time of day is the most important factor for departure delays. On the other hand, Papakostas et al. (2010) pointed out that the flight delay depends on the availability of the resources at each airport. Weather conditions tend to disturb scheduled flight departures and arrivals (Allan et al., 2001; Xu et al., 2008). The airline itself is also a significant factor in departure delays, implying that some airline departures are delayed more frequently than others given the frequent occurrence of late aircraft delays, baggage loading, maintenance, cleaning, and so on (Srivastava, 2011).

All of these factors stemming from the departing airport are reflected in a delayed departure time. Therefore, departure delay time as provided by the departing airport is utilized as a predictor variable. Departure delay time indicates the difference between the actual and scheduled departure time, and is measured in minutes. As with arrival delay times, departure delay times can be negative, indicating an earlier departure than was scheduled. According to panel (a) in Fig. 1, departure delay times and log-transformed arrival delay times seem to have a strong nonlinear positive association. Note that this plot was generated from 2010 data.

In an airborne state, scheduled airborne time (measured in minutes) is an important predictor variable. According to Nikoleris et al. (2012), most flights are able to absorb assigned delays resulting from demand capacity imbalances by reducing their speed. Likewise, departure delayed flights may be able to shorten their delays by increasing their speed. If the scheduled airborne

Table 1
Descriptive statistics for arrival delays (unit: minute) and log-transformed arrival delays.

Variable	Min	Q1	Q2	Q3	Max	Mean	SD	Skew
Arrival delay	−81	−15	−7	5	828	0.009	30.81	4.38
Log-transformed	4.23	4.91	4.96	5.00	5.04	6.89	0.16	2.22

Note: Q1, Q2, and Q3 stand for first, second, and third quartile, respectively.

Table 2
Sources of arrival delays or early arrivals and the measured variables.

Source	Specific factors	Measure variables
Departing airport	Airport capacity - Seasonal factor - Regional factor Weather condition Airline	Departure delay time
Airborne state	Airborne time	Scheduled airborne time
Arriving airport	Airport capacity - Seasonal factor - Regional factor Weather condition Airline	Time of day Day of month Month of year Weather condition Airline

time is relatively long, an aircraft pilot has little room to adjust the arrival time by controlling the flight speed. According to panel (b) in Fig. 1, a slightly negative relationship between the log-transformed arrival delay and scheduled airborne time was significant.

The conditions at the arriving airport are similar to those at the departing airport. Airport capacity may vary because of congestion time, and how busy an airport is depends on a particular time. These factors can be measured using seasonal and regional variables such as scheduled arrival time of day, day of the month, month of the year, and airport city name. However, we consider only one arrival city, and the regional variable is not considered.

Time of day is measured in minutes; for example, if the scheduled arrival time is 13:15, then this time is transformed into 795 min Tu et al. (2008) and AhmadBeygi et al. (2008) argued that delay propagation effects continued for just one day. In particular, Tu et al. (2008) employed time of day into their spline smoothing model to reflect the daily propagation effect. If an intra-day pattern exists in the delay propagation, this variable seems to be reasonable for representing the effect.

Fig. 2(a) is a scatter plot for time of day and log-transformed arrival delays, and shows that identifying a pattern between the two variables is very difficult. In contrast, an evident pattern seems to exist between time of day and log-transformed arrival delays in panel (b) in Fig. 2, where the spline smoothing curve fits to log-transformed arrival delays using the time of day are delineated. According to this figure, delay propagation increases later in the day. Airport congestion may depend on some seasonal variables such as day of month and month of year as well, which will be employed as predictor variables.

Because we consider bad weather conditions, it may initially seem that we would require weather forecasts. However, we applied the actual hourly weather condition data without any weather forecasting. We assumed that no significant difference exists between actual and forecasted weather conditions given the advancements made in the technology for short-term weather forecasting in recent years. In particular, short-term forecasting of a few hours is highly accurate. Because our model uses short-term forecasting, weather information can be collected just after a flight takes off. Further, maximum scheduled airborne times are not longer than 7 h. Moreover, 95 percent of scheduled airborne time is shorter than four and a half hours. In our analysis, bad weather conditions that occur were coded as 1, otherwise 0. The proportion of bad weather events was approximately 10% of the total. Bad weather conditions include rain, snow, fog, ice pellets, haze, drizzle, mist, and thunderstorms.

3. Models and evaluation

3.1. Nonparametric additive model

Let y_i be the arrival delay of flight i of airline (air) with departure delay (dep), which is scheduled to arrive at time (t) on day (d) of month (m). The weather condition (w) at Denver International Airport for the scheduled arrival time was also considered. Let the scheduled airborne time of this flight be written as h. The

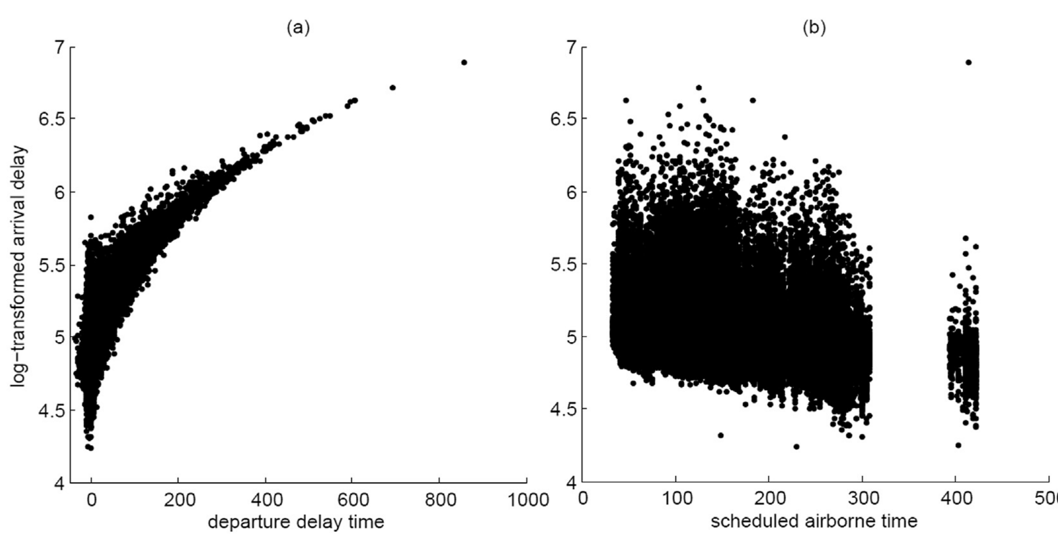

Fig. 1. Scatter plots for (a) departure delays and (b) scheduled airborne time regarding log-transformed arrival delays.

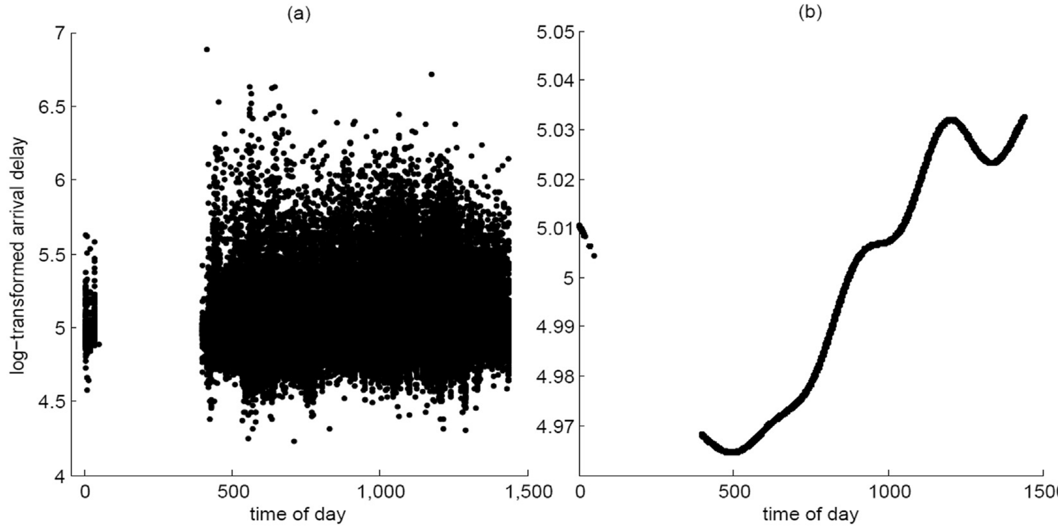

Fig. 2. (a) Scatter plot of time of day and log-transformed arrival delays and (b) spline smoothing curve fits using time of day and log-transformed arrival delays.

considered nonparametric additive models are as follows:

(S1) $y_i = \beta_0 + s(t_i) + s(d_i) + s(m_i) + \varepsilon_i$

(S2) $y_i = \beta_0 + s(t_i) + s(d_i) + s(m_i) + s(dep_i) + \varepsilon_i$

(S3) $$y_i = \beta_0 + s(t_i) + s(d_i) + s(m_i) + s(dep_i) + s(h_i) + \sum_{k=1}^{6} \gamma_k air_{ik} + \kappa_1 w_i + \varepsilon_i,$$

where $s(\cdot)$ is modeled using smoothing splines. In this paper, we implemented the R package "mgcv" (Wood, 2012) for smoothing spline fitting, which is widely applied in the literature. Note that improving forecasting accuracy is possible by tuning parameter sets of smoothing splines.

Regarding the smoothing parameter estimation, literature suggested various types of criteria. This study adopted the generalized cross-validation (GCV) method, which does not require any distribution assumption on the error term. Given this property, we separately modeled the mean function and the error term. As noted in Section 4, the residuals were independently modeled using the skew t distribution.

The predictor variables in S1 consist of only seasonal factors. Here, day of month and month of year are nominal, but this study treats these seasonal factors as numerical variables. As Tu et al. (2008) pointed out, spline smoothing is useful to model the atypical nonlinear form for seasonal and daily delay patterns. Note that Tu et al. (2008) employed only the time of day and day of month factors in predicting departure delays. Additionally, S2 includes departure delay time, a variable obtained after a flight has taken off. As noted, a key finding of this study is that departure delay time is the most important factor for improving short-term forecasting accuracy. Thus, to determine the degree to which it improves prediction performance, we included this variable in S2. S3 additionally includes airborne time, airlines, and weather conditions.

3.2. Benchmark models

Linear regression and median regression were employed as benchmark models. The literature applied linear regression models to flight delay forecasting (Srivastava, 2011; Rebollo and Balakrishnam, 2012). Meanwhile, to the best of our knowledge, previous studies did not suggest the median regression. The median regression is considered because of the skewedness properties of the response variable as notated in Table 1.

The following five linear regression models are implemented:

(L1) $y_i = \beta_0 + \beta_1 t_i + factor(d_i) + factor(m_i) + \varepsilon_i$

(L2) $y_i = \beta_0 + \beta_1 t_i + factor(d_i) + factor(m_i) + \beta_3 dep_i + \varepsilon_i$

(L3) $$y_i = \beta_0 + \beta_1 t_i + factor(d_i) + factor(m_i) + \beta_3 dep_i + \beta_4 h_i + \sum_{k=1}^{6} \gamma_k air_{ik} + \beta_5 w_i + \varepsilon_i$$

We considered three linear regression models, L1, L2, and L3. They have the same predictor variables as in S1, S2, and S3, respectively, but their estimation technique is based on the least squares method. Here, the seasonal variables such as day and month were modeled using dummy variables, which are notated as$factor(\cdot)$. For example, $factor(m)$ can be rewritten as $\sum_{j=1}^{11} w_j \cdot m_{ij}$, where $m_{ij} = 1$ if flight i is scheduled to arrive at month$j$. Otherwise, $m_{ij} = 0$.

The median regression model was applied to forecasting flight arrival time attributable to the pattern of asymmetrically distributed arrival time and its residuals. When the distribution is skewed to the right, the mean tends to be larger than the median. Thus, the linear regression model using the least squares approach may lead to biased estimation results under the asymmetrically distributed error term. The median regression model may reduce this problem, and can be estimated using quantile functions. Quantile functions were estimated by providing different weights to positive and negative residuals in their optimization scheme (Koenker and Hallock, 2001). This study implemented the median regression model and evaluated the performances of three forecasting models: M1, M2, and M3. These models have the same forms as L1, L2, and L3, respectively, but their estimation technique is based on the median regression scheme.

3.3. Prediction evaluation

The models were estimated using data from 2010, and the forecasting was obtained using data from 2011. As was explained, the original arrival delay N_i is transformed into $y_i = \ln(N_i + 150)$. Estimation and forecasting will be based on y_i, but the prediction evaluation will be based on N_i. For the case of evaluating the arrival time forecast, the following well-known root mean square error (RMSE) and the mean absolute deviation (MAD)-based evaluation criterion is implemented:

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(N_i - \widehat{N}_i\right)^2} \text{and} \quad \frac{1}{n}\sum_{i=1}^{n}\left|N_i - \widehat{N}_i\right|$$

where $N_i = \exp(y_i) - 150$ and $\widehat{N}_i = \exp(\widehat{y}_i) - 150$, respectively. n is the number of forecasted values and $\widehat{y}_i$ is forecasted log-transformed arrival delays.

4. Empirical results

4.1. Prediction results

The estimation was obtained by fitting the 2010 data to the model, where most predictor variables were significant. The multi-collinearity problem was checked by computing the variance inflation factors, which had values of less than 2 for all of the incorporated predictors, indicating that no multi-collinearity exists.

Using each fitted model, we made a prediction regarding the data for 2011. Table 3 summarizes the prediction results among the models and reports the RMSE and MAD. S3 seems to be the most accurate prediction model in terms of both the RMSE and the MAD criterion; these values are shown in bold in Table 3. In summary, the spline smoothing regression model dominantly outperformed the linear regression and median regression models. Median regression showed the most inferior prediction performance among three types of models.

Regarding S-typed models, if additional predictor variables were included, their prediction performances improved. In particular, if the departure delay time is additionally included (S2), the prediction accuracy appears improved from S1. S2 significantly improved from S1 by approximately 60% in RMSE and 47% in MAD. An interesting result is that the departure delay time seems to play a key role in improving prediction performance.

Following a referee's suggestion, we further examined the forecasting performance of models regarding the short term (from January 1, 2011 to June 30, 2011) and the long term (from July 1, 2011 to December 31, 2011). We only focused on S-typed and L-typed models because M-typed models were inferior to others throughout the entire 2011 analysis. The prediction results are summarized in Table 4.

According to this table, S-typed models outperform L-typed models. In particular, S3 is superior to all other models in both short-term and long-term forecasting. Interestingly, the long-term forecasting results seem slightly more accurate than the short-term results in terms of RMSE and MAD. Regardless of this outcome, we need to mention that it may be difficult to claim any significant superiority of prediction performance of the models between the short term and the long term given the different sample sizes (notated in the parenthesis) regarding the two periods.

Lastly, in addition to the aggregate results in Table 3, we considered illustrating the actual delays and forecasting performances using S3 generated by seven airlines and 12 months of a year as case studies. In particular, their box plots are depicted in Fig. 3, which indicates various different distributions of actual delays and forecasts through different airlines and different months of a year. Distributions of actual delays and forecasts appear to be matched at large. According to Fig. 3, some airlines tended to cause more extreme delays than others. Similar patterns were observed in different months of the year.

Table 3
Prediction results for each model using RMSE and MAD for the arrival delays.

Model	RMSE	MAD	Model	RMSE	MAD	Model	RMSE	MAD
S1	31.477	17.118	L1	31.520	17.173	M1	32.358	16.816
S2	12.677	9.071	L2	31.391	10.041	M2	42.623	10.168
S3	**12.200**	**8.631**	L3	30.529	9.638	M3	43.312	9.801

Table 4
Prediction results for S-typed and L-typed models on short-term and long-term forecasting.

Horizon	Model	RMSE	MAD	Model	RMSE	MAD
2011	S1	32.709	17.805	L1	32.760	17.872
short-term	S2	12.690	9.115	L2	37.695	10.168
(87,180)	S3	**12.364**	**8.756**	L3	36.659	9.800
2011	S1	30.159	16.412	L1	30.192	16.454
long-term	S2	12.664	9.026	L2	23.187	9.910
(84,791)	S3	**12.029**	**8.503**	L3	22.550	9.472

4.2. Validity of skew t distribution for error terms

This study attempts to find an appropriate distribution to the residuals and compute the probability of delays using the fitted residual distribution. In particular, we suggest modeling the residuals using the skew t distribution. The residuals were examined after fitting the arrival delay data using nonparametric additive models, which outperformed other models as noted in the prediction result section. As a result, the distributions of the residuals were skewed to the right, indicating the possible appropriateness of applying skewed distributions to model the residuals. Therefore, we attempted several skewed distributions including the skewed normal (Azzalini and Capitanio, 1999) and skew t (Azzalini and Capitanio, 2003) to the model fitting using Kolmogorov−Smirnov (K−S) tests. Among these distributions, the skew t distribution for the fitted residuals of several linear regression models showed acceptable results, suggesting the possibility of modeling the residuals using the skew t distribution. The residuals for the 2010 data were implemented, and Table 4 summarizes the results of the tests. The R package “sn” was implemented for the model estimation (Azzalini, 2012), and the parameter estimates regarding the skew t distribution resulted. Using the estimated parameters, the p-values for the K−S tests were calculated. According to the p-values for the K−S tests, S2 and S3 showed acceptably good fits to the skew t distribution. In contrast, S1 showed unacceptable fits.

Next, we checked whether the two distributions of fitted residuals (the 2010 and 2011 data) are statistically identical. Note that we are appropriately able to apply the estimated distribution from the 2010 data to the computation of the probability of arrival delays for the 2011 data only when the two distributions of the fitted residuals are similar. Therefore, equality between two distributions should be investigated. To check the similarity of the two distributions, the sample quantiles of the two distributions were compared. Using the difference between the two groups of percentiles, we carried out a paired t test. According to the reported p-values for the paired t test in Table 5, the residuals between the years 2010 and 2011 showed no significant difference in all S1−S3.

5. Computing the probability of arrival time

Landry et al. (2013) suggested a method for dynamic taxiway and runway conflict prevention. According to them, in order to avoid conflicts, acceptable number of aircrafts on a taxiway section, a runway section, and an intersection at a given time should be

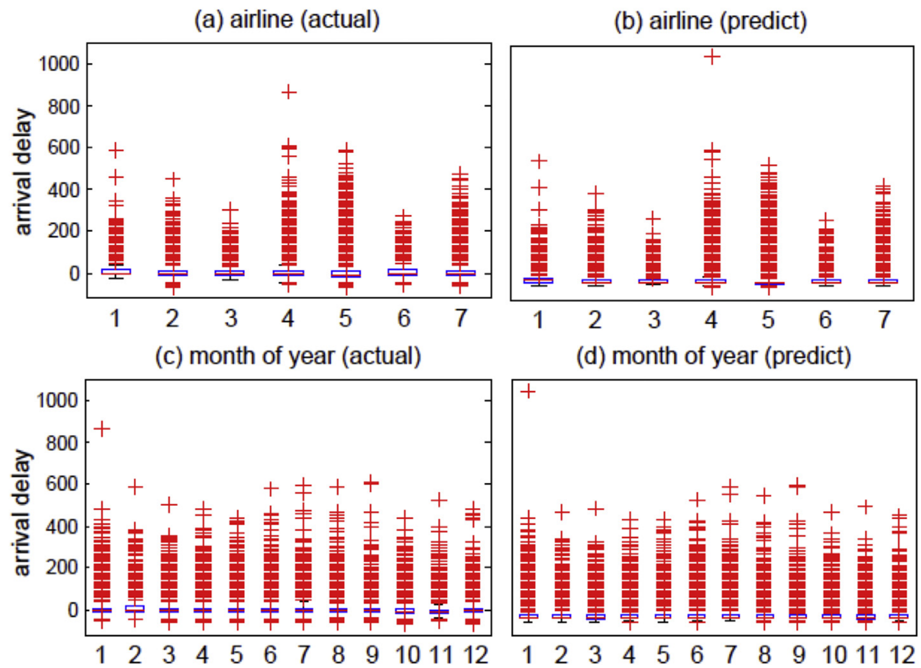

Fig. 3. (a) and (b) are box plots of actual arrival delays and forecasts using S3 for seven airlines, respectively. (c) and (d) are box plots of actual arrival delays and forecasts using S3 for 12 months, respectively.

Table 5
Test results for the validity of the skew t distribution and two density equalities regarding the fitted residuals.

Fitted	K–S test	Skew t				Paired t
Model	p-value	Location	Scale	Shape	d.f.	p-value
S1	0.000	−0.106	0.116	1.383	2.935	0.431
S2	0.902	−0.011	0.066	0.179	4.753	0.866
S3	0.792	−0.038	0.066	0.751	4.829	0.876

Note: d.f. indicates degree of freedom.

specified and constrained. For appropriate time capacity constraints, accurate time prediction of aircraft arrival for each sector seems to be important. On the other hand, the Volpe National Transportation Systems Center operates the enhanced traffic management system (ETMS) to handle air traffic flows in the United States. The ETMS generates an alert if traffic demand exceeds a sector's or an airport's capacity. The level of traffic demand is measured by an alert threshold. For example, alert thresholds for a sector indicate the number of aircrafts in the sector at a point in time.

Tu et al. (2008) applied the concept of alert threshold for a sector to provide the probability of departure delays. This study modifies and applies their approach to the case of arrival time. The methodology to compute the probability of an arrival delay is similar to that for a departure delay in Tu et al. (2008), except that our approach focuses on arrival time and skew t distribution is implemented. Using similar representation and the example in, Tu et al. (2008), we suggest a method for computing the probability of arrival time as follows. Let $I_i(s,t)$ be 1 if a flight i is in sector s at time t; otherwise, 0. $N(s,t)$ is the number of flights in sector s at time t. The following equalities then hold:

$$\mathrm{E}[N(s,\ t)] = \sum_{all\ \ i} \mathrm{E}[I_i(s,\ t)] \ \text{ and } \ \mathrm{E}[I_i(s,\ t)] = P[I_i(s,\ t) = 1]. \tag{1}$$

Assume that a given flight has not yet arrived. We are then able to estimate the probability of the arrival delay of flight i for a given time t. Let Δ denote the required time for a flight i to enter and exit the sector s from the current flight plan. Note that sector s is the sector in which the flight is expected to arrive.

$$P[I_i(s,\ t) = 1] = P\left(t - \Delta \le t^i_{arr} \le t + \Delta\right), \tag{2}$$

where $t^i_{arr} = t^i_{sch} + N_i$ and $\ln(N_i + 150) = y_i = f_i + \varepsilon_i$, or $N_i = \exp(f_i + \varepsilon_i) - 150$. The flight arrival time t^i_{arr} is composed of the scheduled arrival time plus arrival delay or early arrival. The log-transformed arrival delays can be decomposed into the mean function f_i and error term ε_i.

This approach is applied to calculate the probability of arrival delay. Assume that the present time is later than the scheduled arrival time and that the flight has not yet arrived at the present time. For example, let flight i be scheduled to arrive at 10:00 a.m. on March 5. Suppose that Δ is 10 min. We want to compute the probability that this flight arrival will be delayed at given time t of 10:20 a.m. using this information.

f_i can be estimated using the assumed model and predictor information (*dep, h, t, d, m, air1–air6*) for flight i. For example, we consider model S3, whose estimated coefficients are utilized to

$$\begin{aligned}
&P\left(10:10 \le t^i_{arr} \le 10:30\right)\\
&= P\left(10:10 \le t^i_{sch} + N_i \le 10:30\right) = P\left(10:10 \le t^i_{sch} + \exp(f_i + \varepsilon_i) - 150 \le 10:30\right)\\
&= P(160 \le \exp(f_i + \varepsilon_i) \le 180) = P(\ln(160) \le f_i + \varepsilon_i \le \ln(180))\\
&= P(5.075 \le f_i + \varepsilon_i \le 5.193)
\end{aligned} \tag{3}$$

obtain the estimates of f_i under the predictor information of flight i. Meanwhile, the variable ε is known to have a skew t distribution with the parameters (location: −0.038, scale: 0.066, shape: 0.751, degree of freedom: 4.829) of S3 in Table 5.

If the estimate of f_i is 5 min, then Eq. (3) is rewritten as:

$$P(0.075 \le \varepsilon_i \le 0.193) = P(\varepsilon_i \le 0.193) - P(\varepsilon_i \le 0.075) \tag{4}$$

The right part of Eq. (4) is obtained by computing the cumulative distribution function of the skew t distribution. If we compute the expected number of flights in sector s at time t, $\mathrm{E}[N(s,t)]$ in Eq. (1) is utilized. If we compute the accurate expected number of flights at some designated zone, more effective traffic control by aviation systems will be available. Finally, the arrival delay rate may be reduced.

6. Conclusion and discussion

This article suggested applying spline smoothing-based nonparametric additive models to the short-term forecasting of flight arrival times at Denver International Airport during 2011. Some information from the time at which a flight takes off from a departing airport was employed in the forecasting. Departure delay time, scheduled airborne time, airlines, and weather conditions were shown to be important to improving prediction accuracy, along with the seasonal variables. Among these, departure delay significantly improved forecasting accuracy. We also suggested a method to compute the probability of the arrival time of a flight by fitting the skew t distribution to the residuals of the model under consideration, which helps compute the accurate expected number of arriving flights.

As it was mentioned in the introduction section, airline delay costs have been tremendous. This type of cost is an additional to the regular operation cost. If we have more accurate prediction for flight arrival times, we will be able to increase the chances of reducing the related delay costs. For example, we can build more appropriate flight scheduling, maintenance planning, and so on. These activities will improve the qualities of air transportation management.

Forecasting accuracy using S3 model improved by about 60% and 10% in terms of RMSE and MAD, respectively, compared with L3. Although the precise quantitative impact of improved accuracy of arrival delays to the air transportation management was not rigorously investigated in this study, it will lead to a significant result. Those studies will be considered in the future.

To conclude, there are limitations to this research. We did not consider security delay as a predictor variable for arrival delays in this paper. Although the share of security delay in the causes of flight delays is negligible, its influence is not negligible. Therefore, a systematic approach to determine its pattern should be developed and reflected in the model, which will help improve forecasting accuracy. We expect to provide these trials in future research.

Acknowledgment

I am very grateful to the editor and two anonymous referees for their helpful comments and suggestions. This work has been supported by a Sogang University Research Grant (201310017.01).

References

AhmadBeygi, S., Cohn, A., Guan, Y., Belobaba, P., 2008. Analysis of the potential for delay propagation in passenger airline networks. J. Air Transp. Manag. 14, 221–236.

Allan, S.S., Gaddy, S.G., Evans, J.E., 2001. Delay Causality and Reduction at the New York City Airport Using Terminal Weather Information Systems. Lincoln Laboratory Massachusetts Institute of Technology Lexington, Massachusetts, 2001.

Azzalini, A., 2012. The Skew-normal and Skew-t Distributions. R Package.

Azzalini, A., Capitanio, A., 1999. Statistical applications of the multivariate skew-normal distribution. J. R. Stat. Soc. Ser. B 61, 579–602.

Azzalini, A., Capitanio, A., 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. J. R. Stat. Soc. Ser. B 65, 367–389.

Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., Peterson, E., Sherry, L., Trani, A., Zou, B., 2010. Total Delay Impact Study. A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States. The National Center of Excellence for Aviation Operations Research. http://www.isr.umd.edu/NEXTOR/pubs/TDI_Report_Final_10_18_10_V3.pdf.

Deshpande, V., Arıkan, M., 2012. The impact of airline flight schedules on flight delays. Manuf. Serv. Operations Manag. 14, 423–440.

Idris, H., Clarke, J.P., Bhuva, R., Kang, L., 2002. Queuing model for taxi-out time estimation. Air Traffic Control Q. 10, 1–22.

Koenker, R., Hallock, K.F., 2001. Quantile regression. J. Econ. Perspect. 15, 143–156.

Landry, S.J., Chen, X.W., Nof, S.Y., 2013. A decision support methodology for dynamic taxiway and runway conflict prevention. Decis. Support Syst. 55, 165–174.

Mayer, C., Sinai, T., 2003. Network effects, congestion externalities, and air traffic delays: or why not all delays are evil. Am. Econ. Rev. 93 (4), 1194–1215.

Mueller, E.R., Chatterji, G.B., 2002. Analysis of aircraft arrival and departure delay characteristics. In: Proc. AIAA Aircraft Tech., Integration, and Oper. Forum at Aviation Week's Aerospace Expo (Los Angeles).

Nikoleris, T., Chatterji, G.B., Almog, N., Palopo, K., 2012. Arrival delay absorption using extended metering with speed control. In: 12th AIAA Aviation Technology, Integration, and Operations Conference and 14th AIAA/ISSM (Indianapolis, Indiana).

Papakostas, N., Papachatzakis, P., Xanthakis, V., Mourtzis, D., Chryssolouris, G., 2010. An approach to operational aircraft maintenance planning. Decis. Support Syst. 48, 604–612.

Rebollo, J.J., Balakrishnam, H., 2012. A network-based model for predicting air traffic delays. In: 5th International Conference on Research in Air Transportation (ICRAT).

Schumer CE and Maloney CB (2008). Your flight has been delayed again: Flight delays cost passengers, airlines, and the U.S. economy billions. Report, Joint Economics Committee Majority Staff of U.S. Senate, chaired by Senator Charles E. Schumer and vice-chaired by Rep. Carolyn B. Maloney.

Srivastava, A., 2011. Improving departure taxi time predictions using ASDE-X surveillance data. In: Digital Avionics Systems Conference (DASC), IEEE/AIAA 30th.

Tu, Y., Ball, M.O., Jank, W.S., 2008. Estimating flight departure delay distributions-A statistical approach with long-term trend and short-term pattern. J. Am. Stat. Assoc. 103, 112–125.

Wood, S.N., 2012. Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation. R Package.

Xu, N., Sherry, L., Laskey, K.B., 2008. Multifactor model for predicting delays at U.S. Airports. Transp. Res. Rec. J. Transp. Res. Board 2052, 62–71.

Myung Suk Kim is a Professor in School of Business at Sogang University, Seoul, South Korea. He received MS and Ph.D. in statistics from Texas A&M University. Before joining Sogang University, he worked as a time series model analyst at JP Morgan Chase & Company, Columbus, Ohio, USA. His research interests include forecasting (energy, flight arrival, etc.), risk measurement, nonparametric model validation, business data modeling (CRM), and environmental statistics. He has published many papers in numerous well established journals such as European Journal of Operational Research, Computational Statistics & Data Analysis, Expert Systems with Applications, Industrial Management & Data Systems, Journal of Risk Model validation, Environmetrics, and so on.