Airline Delay Time Series Differentials:
Autoregressive Integrated Moving Average Model

Ronald Wesonga, Makerere University, Kampala, Uganda

Fabian Nabugoomu, Kyambogo University, Kampala, Uganda

Brian Masimbi, Makerere University, Kampala, Uganda

ABSTRACT

Flight delays affect passenger travel satisfaction and increase airline costs. The authors explore airline differences with a focus on their delays based on autoregressive integrated moving averages. Aviation daily data were used in the analysis and model development. Time series modelling for six airlines was done to predict delays as a function of airport's timeliness performance. Findings show differences in the time series prediction models by airline. Differential analysis in the time series prediction models for airline delay suggests variations in airline efficiencies though at the same airport. The differences could be attributed to different management styles in the countries where the airlines originate. Thus, to improve airport timeliness performance, the study recommends airline disaggregated studies to explore the dynamics attributable to determinants of airline unique characteristics.

Keywords: *Airline, ARIMA Model, Delay Differentials, Developing Airport, Forecasting*

INTRODUCTION

Most of the foreign exchange earnings of growing economies such as Uganda largely depend on transactions conducted through airports (Voltes-Dorta & Lei, 2014; Zhang, Yang, Wang, & Zhang, 2014). Efficient operation at such airports is therefore mandatory to nurture a strategic position towards economic development of the country. Air traffic flow at Entebbe International Airport, the only international airport in Uganda, has greatly increased since the year 1991, faster than was projected to grow. Suffice to note that international passenger numbers sprung from a mere 118,000 in 1991 to over one million by 2011. Such an increase often relates to more

DOI: 10.4018/IJASOT.2014070105

congestion and delays (Bilotkach & Lakew, 2014; Civil Aviation Authority, 2012; Lam, Tang, Chan, & Tam, 2006; Noland, Quddus, & Ochieng, 2008; Zou & Hansen, 2014). Flight delays do not only affect passenger satisfaction but also bring along costly consequences to airlines.

According to the Civil Aviation Authority (2012) in the year 2007, sixteen international airlines had scheduled operations to and from the airport, serving fourteen different destinations. Consequently, EIA has developed as Uganda's international gateway with its traffic growing from a paltry 27,000 passengers in 1962 to 1.2 million at the turn of Uganda's fiftieth independence anniversary. The growth in traffic has inadvertently led to significant delays in air traffic flow at the airport as often predicted (Ryerson, Hansen, & Bonn, 2014).

As a standard, a flight is considered delayed if it leaves the gate more than fifteen minutes from the originally scheduled departure time (International Civil Aviation Organization, 1995; Kanafani & Ghobrial, 1985). However, in the context of this study, the interest was on the total expected additional time that should be included for departing flights from the airport after the scheduled time of departure.

A summary of the different methods for analysing delays was made into five categories: regression and related methods, time series analysis, Bayesian networks model analysis, and cluster and classification analysis and simulation (Konishi & Kitagawa, 2007; Ninj, 2007). It was further noted that regression included methods for using observations to predict or explain the delays. In time series analysis, the trend analysis, spectral analysis, and Markov Chain analysis are introduced. Regression analysis approach is used in the analysis of arrival delays given different causes of the delays (Gano & Banavar, 2005; Sridhar, Grabbe, & Mukherjee, 2008).

In this paper, non-traffic flow delays at airports have been correlated with sets of causal factors and created models to predict aggregate delays at airports on a daily basis. In order for this study to be consistent with the way traffic is managed, Wesonga and Nabugoomu (2014) and Wesonga, Nabugoomu, and Jehopio (2012) proposed and evaluated models of causal factors of delays that would provide the analytical bases for contributing toward the improvement of the efficiency for traffic flow management at Entebbe International Airport (EIA). In their analysis, parameterized approaches were applied to develop models that determine airport delay where both aviation and meteorological parameters have been studied (Zou & Hansen, 2014).

DATA AND METHODS

Data Management

To achieve objectives of the study, a number of tools were applied to the data collected. Aviation data, collected for the years 2006, 2007, and 2008 included daily records for all twelve months of the year. The dataset had eight attributes, namely; date, operator (airline/carrier), type of aircraft, nationality (origin), from/to (where the aircraft came from and its destination), category (international/local), expected time of arrival/departure (ETA/D), and actual time of arrival/departure. Samples of six airlines considered to have busy schedules were identified: Eagle Air (EA), Kenya Airways (KQ), South African Airways (SAA), Royal Dutch Airlines (KLM), Ethiopian Airlines (ET), and British Airways (BA). To achieve the stated objectives and hypotheses, four relevant fields were extracted: 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 formula in Equation 1 was used in the computation of the delays for each record:

$$Delay_i\left(minutes\right) = ATA / D_i - ETA / D_i$$

where i (i = 1, 2, 3…N) represents an airline or aircraft.

Tests of Normality for Airline Delay

The test of normality for the airline delay was deemed important since having normally distributed errors is equivalent to having normally distributed airline delay for any linear time series model. Normality assumption was made in the development of airline time series model for the purpose of ensuring that results are generalizable and inferences about the population parameters can easily be made. Furthermore, normality of errors of estimation is often assumed in using the Akaike Information Criteria (AIC) for order selection and in computing prediction intervals. When using maximum likelihood estimation (MLE), a distributional assumption is applied, in this case probability normal distribution (Gaussian). So, if the Gaussian assumption does not hold, then the likelihood function is messed up and cannot be reliable.

The Shapiro-Wilk test produced the following analysis (W = 0.63388, P-value = 0.0000, N = 504) confirming that the delays were not normally distributed. A natural log transformation was therefore done on the delay variable and the Shapiro-Wilk test produced the following test (W = 0.98754, P-value = 0.00055, N = 462), implying that the log transformations were normally distributed. This result has two implications: the estimation errors are normally distributed or follow the Gaussian distribution and that the sample data is good enough to be used for drawing inferences about airline delays at the airport. Therefore, more reliable airline delay prediction can be made with confidence.

Three other variables were introduced as numerical representatives: day of the week, month, and airlines:

1. **Day:** 1-Monday, 2-Tuesday, 3-Wednsday, 4-Thursday, 5-Friday, 6-Saturday, 7-Sunday;
2. **Airline:** 1-Eagle Air, 2-British Airways, 3-Kenya Airways, 4-SAA, 5-Ethiopian Airlines, 6-KLM;
3. **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.

Therefore, the final dataset had the following variables namely, dates (day of the week, date, month, and year); arrival and departure times (actual and scheduled); airlines: Eagle Air, KQ, SAA, BA, Ethiopian Airlines, and KLM. Thus the derived variables for the study dataset included average delays (natural logarithms), month, day and Airline, schedule per day. These variables come from the study dataset that was collected at Entebbe International Airport. They help in the objective of this study to explore the characteristics associated with airline delay and develop appropriate airline-based time series models.

Data Analysis

We assume the airline delay Y_t at time t, representing day. To model using an autoregressive integrated moving average (ARIMA) (p, d, q), we require the AR (p), degree of differencing d and moving average, MA (q).

Descriptive time series analysis that includes numerical summaries was used to show the overall distribution of the airlines' delays at the case study. Understanding the general characteristics of airlines was important for providing the initially required information for developing time series models. Aggregation enabled us to view the series at coarser intervals (quarterly and monthly basis) and also to see the difference between the quarterly and monthly series. The comparison enabled us to determine whether there was a systematic difference between the airlines. The two airlines chosen for comparison were based on the distance each travels. That is, BA was compared with KLM, SAA with Ethiopian Airways, and Eagle Air with KQ. Airline timeliness performance is often associated with distance travelled from the origin (EIA). The more similar the distances travelled, the more likely that the correlation between airline delays will be strong. In performing pairwise comparisons, the study was guided by a number of supporting reasons for a more profound understanding of the behaviour of airline operational performance. The paired airlines tend to be clustered around many other similar characteristics such as planned costs (the cost of an airline plan under the assumption that all flights occur as scheduled and without disruption), time of flight or flight duration, ground delay programs, customer behaviour, routing behaviour, and airline management styles.

Modelling Airline Delay Time Series

ARIMA models provide an approach to time series forecasting. Exponential smoothing and ARIMA models are the two most widely-used approaches to time series forecasting, and provide complementary approaches to the problem. While exponential smoothing models are based on a description of trend and seasonality in the data, ARIMA models aim to describe the autocorrelations in the data and also offer ways of integrating with autoregressive and moving averages in the data.

In time series analysis, the ARIMA model is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data to help better understand the data and also to predict future points in the series. They are applied in some cases where data show evidence of non-stationarity, where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to remove the non-stationarity.

The general form of the model is ARIMA (p, d, q), where parameters p, d, and q are non-negative integers that refer to the order of autoregressive, integrated, and moving average parts respectively. ARIMA models form an important part of the Box-Jenkins approach to time-series modelling.

The formulation had a model form and several parameters using the approach to time series analysis, ARIMA for each airline (Nihan & Holmesland, 1980). In particular, ARIMA models were developed for airline delay time series to help better understand the different airline delays at the airport and also to develop models for predicting airline-based delays. We focus on the ARIMA model due to the mainly non-stationary property of the data collected.

Algorithm in Time Series Model Building Process Using ARIMA

The study applied the following three-stage procedure of time series modelling approach, namely model identification, parameter estimating, and diagnostics for model suitability.

Step One

This step comprised of identification of the model and included checking whether the time series data was stationary by determining if any differencing was needed:

1. Two indications that a series is not stationary, namely, a time series that appears to have different overall levels or degrees of autocorrelation in different sections of the series and autocorrelation function (ACF) using Correllogram that does not decay to zero;
2. Model identification was then performed to examine the ACF and partial autocorrelations (PACF). If the process was AR (*p*), the partial autocorrelation was zero at lag $\geq p + 1$. So, the sample partial autocorrelation plot was examined to identify the order. If the process was MA (q), the autocorrelation was zero at lag $\geq q + 1$. So, the sample autocorrelation plot was examined to identify the order.

Step Two

Given that the time series models fitted in this study are of the form ARIMA (p, d, q), estimation of the model parameters was done using the popular ARIMA () function in the R statistical software. This was supplied with three parameters, which were used to specify the model type, conventionally known as *p* (the autoregressive (AR) order), *d* (the degree of differencing of the airline delay), and *q* (the moving average [MA] order). Obtaining ARIMA (p, d, q) model parameters p, d, and q help in developing a more precise model that forecasts airline delay at minimal error of estimation. Therefore, this step plays a very significant role in our modelling process of the airline delays. It is this step that results into the desired ARIMA (p, d, q) models for the six airlines under study.

Step Three

This step deals with model checking, an important function in statistics; the step ensures that the correct model is well fitted to the data. Many different models seem to be appropriate, but only one is preferred, hence the search and therefore the model checking step. For example, information criteria such as the Akaike Information Criteria (AIC) or Bayesian Information Criteria (BIC) could be used for model checking. These basically quantify the goodness of fit and the simplicity or parsimony of the model into a single statistic such that when comparing two models, the one with the lower AIC is generally preferred. For the case of our study, model checking was done through three diagnostic plots for ARIMA model. These included standardized residuals or error terms that should show no pattern with time, ACF of residuals that should have no significant autocorrelations, the Ljung-Box statistic for the null hypothesis of independence in the time series of residuals.

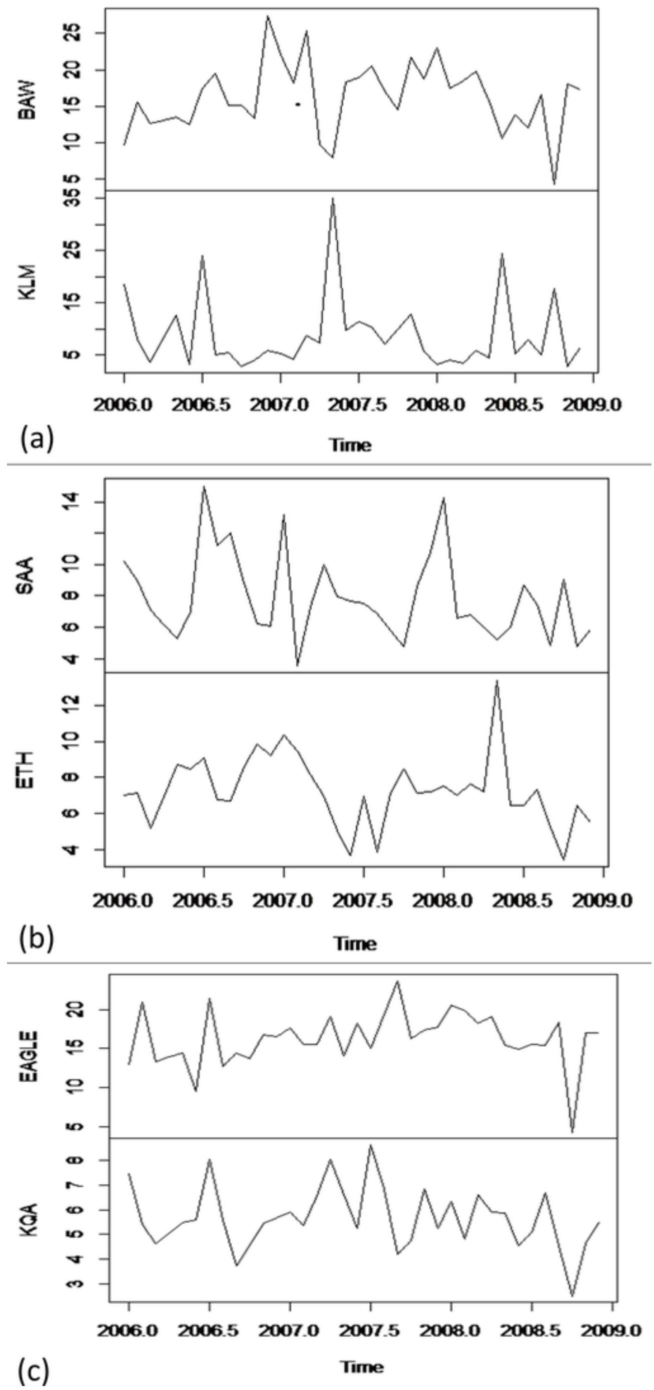
Residuals are errors made due to estimation. In time series analysis using ARIMA modelling approach, we check the residuals to see if there are no patterns and establish if they are really random or whether they are influenced by other occurrences. The objective here was to minimise any error due bias as much as possible. On the other hand, even within the residuals, the correlations between them are not expected to be high; otherwise, this would indicate existence of bias in model estimation.

FINDINGS

Comparison of the Time Series of the Different Airlines

In Figure 1, we present a comparison of delays by pairs of airlines. Comparisons were done only between airlines that cover approximately the same distances from EIA, taken as the origin. As explained under data analysis section, the paired airlines tend to be clustered around many

Figure 1. Pairwise comparison of airline delay time series based on distances covered: (a) Comparison of time series of BA and KLM; (b) Comparison of time series South African Airline and Ethiopian Airlines; (c) Comparison of time series of Eagle Air and Kenya Airways



similar characteristics. Therefore, to be able to reduce the time series model estimation errors due to biases caused by differences in distance covered, pairwise airline delay analyses were done between airlines that traverse approximately equal distances from Entebbe International Airport.

Figure 1 (a) shows that KLM had delays which are closer to zero and also to each other as compared to BA, but it was further noted that BA appeared to have more rapid fluctuations than KLM. While Figure 1 (b) showed that Ethiopian Airlines had delays which are closer to zero and also to each other as compared to SAA, but SAA appeared to have more rapid fluctuations.

Comparisons of delay between airlines based on the length of distances often travelled.

Comparisons are made between BA and KLM (approx. 10,000 km); ETH and SAA (approx. 5,000 km) and KQ and EAGLE (approx. 900 km). These pairs of airlines cover approximately equal distances as approximated and shown in brackets.

Figure 1 (c) showed that KQ had delays which were closer to zero, but Eagles Airways had delays which were closer to each other. Thus, it was noted that KQ appeared to have more rapid fluctuations.

Checking for Stationarity of the Airline Delays

Given the objective of exploring the ARIMA time series models for each of the six airlines operating at Entebbe International Airport, we had to ensure that the airlines' delays are stationary. Checks for stationarity were done through plots of ACFs against lags for each airline and results showed that the original time series for KLM, SAA, Ethiopian Airlines, and Eagle Air were stationary, implying that the difference-value (d) for delays of these airlines was equal to zero (d = 0). However, from analysis, the original time series for delay by two airlines; BA and KQ were not stationary. However, checking their first differences showed that they were stationary. Given that we did not have any further evidence against the difference for the two airlines being equal to one, we then defined the time series models as ARIMA (p, d = 1, q) for BA and KQ airlines respectively.

Determining the AR and MA Degrees

The autoregressive (AR) and moving average (MA) degrees p and q refer to the maximum values beyond which the effect of past values of the airline delays will not cause significant effect on the current airline delays. These are often written as AR (p) and MA (q) and are important in ARIMA (p, d, q) model formulation for each airline.

Analyses of the ACFs and the PACFs by airline were done to show the days where significance occurred. This process was repeated for all airlines in an attempt to fit appropriate ARIMA models. Partial autocorrelation plots are a commonly used tool for identifying the order of an autoregressive model (Box, Jenkins, & Reinsel, 2013). Since it is known that the partial autocorrelation of an AR (p) process is zero at lag p + 1 and greater. Therefore, for our case, since the sample autocorrelation indicated that an AR model would be appropriate, the sample partial autocorrelation was examined to help identify the order. We looked for the point where the partial autocorrelations for all higher lags are essentially zero. Placing an indication of the sampling uncertainty of the sample PACF was helpful for this purpose because this is usually constructed on the basis that the true value of the PACF at any given positive lag is zero.

Further scrutiny revealed that there was no significant autocorrelation and partial autocorrelation for the KLM airline. Therefore, we defined that AR (0) and MA (0), both had zero degrees. Since p = 0, d = 0 and q = 0, then the model for the KLM airline was established as simply white noise model, denoted by the ARIMA (0, 0, 0) model.

For the case of British Airways, it was observed that there were significant autocorrelations at the following days: 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 19, 21, 22, and 27; and significant partial autocorrelation at days 1, 2, 3, 4, 5, 6, 8, 9, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, and 31 (note that "lag" on the our plots refers to the 31-day cycle). We concluded that the AR had degrees up to 31 lags and MA had degrees up to 27 lags. But when small lags were fitted in model one at a time, it was established that $p = 5$, $d = 1$ and $q = 3$. Thus, the time series model for the British Airways, BA airline is denoted by the ARIMA (5, 1, 3) model. Thus, the autoregressive degree for BA was AR(5) and the moving average was MA(3) implying that up to five lags of previous delays for BA would still significantly impact on the current airline delays. On the other hand, only three degrees of the moving averages would cause variations in the current BA delay. When more historic datasets for airline delay significantly cause changes in the current values, two main implications would be derived: airline system sustainability and level of quality management style implemented over time.

SAA showed significant autocorrelation at 2, 5, 7, 9, 10, 12, and 30 days and significant partial autocorrelation were at 2, 5, 7, 9, 12, and 14 days. The study concluded that the AR had degrees up to 14 lags and MA had degrees up to 30 lags. When we fitted small lags in the model one at a time, it was established that $p = 5$, $d = 0$, and $q = 2$. Thus, the model for the SAA airline is denoted by the ARIMA (5, 0, 2) model. Like BA, SAA's autoregressive degree was AR(5) and the moving average given by MA(2) implying that up to five lags of previous delays for SAA would still significantly impact on the current airline delays. On the other hand, only two degrees of the moving averages are found to cause variations in the current SAA delay.

Similarly, for the case of Ethiopian Airlines (ETA), the significant autocorrelation were observed at days 7 and there was no significant partial autocorrelation. It was thus concluded that the AR had degrees equal to zero and MA was 7. Therefore, the ARIMA model arguments are $p = 0$, $d = 0$ and $q = 7$, implying that the model for Ethiopian Airlines is denoted by the ARIMA (0, 0, 7) model. Unlike BA and SAA, Ethiopian Airlines' ETA could be predicted using the moving average of order seven MA(7), implying that up to seven degrees of the moving averages would cause variations in the current ETA delay.

The significant autocorrelation for KQ were at days 1 to 23, 27, and 29; and significant partial autocorrelation were observed at days 1 to 24, 29, 30, and 31. Therefore, it was concluded that the AR had degrees up to 31 lags and MA had degrees up to 29 lags. But by fitting small lags in the model, one at a time, we established that $p = 4$, $d = 1$, and $q = 3$. Thus, the model of KQ airline is denoted by the ARIMA (4, 1, 3) model. KQ on the other hand required autoregressive degree of order four AR(4) and the moving average of order three MA(3); implying that up to four lags of previous delays for KQ would significantly impact on the current airline delays. Relatedly, three degrees of the moving averages were found to cause variations in the current KQ delay.

Lastly, for Eagle Air, results show that the significant autocorrelation were at days 1, 2, 4, 5, 15, 18, and 31; while significant partial autocorrelation were at days 1, 3, 13, 26, and 29 (where lags on the plots used the 31-day cycle). Thus, we concluded that the AR had degrees up to 29 lags and MA had degrees up to 31 lags. On fitting small lags in the model, one at a time, it was established that $p = 3$, $d = 0$, and $q = 2$. Thus the derived model of Eagle Air (EA) airline is denoted by ARIMA (3, 0, 2) model. EA required autoregressive degree of order four AR(3) and the moving average of order two MA(2); implying that up to three lags of previous delays for EA would significantly impact on the current airline delays. Two degrees of the moving averages were found to cause variations in the current EA delay.

Deriving the ARIMA Models for the Airlines (KLM, BA, SAA, ET, KQ, EA)

In Table 1, we present the derived actual ARIMA time series models for the different airlines operating at Entebbe International Airport. These models can be used for prediction of airline-based delays at the airport.

ARIMA Model Check by Airline

Model checking was done for each airline using diagnostic plots for standardised residuals, ACFs of residual, and p-values for Ljung-Box statistics. From definition, the Ljung-Box test statistic gets larger as the sample auto-correlations of the residuals get larger. Its p-value is the probability of getting a value as large as or larger than that observed under the null hypothesis, which are that the true residuals for the estimated airline delays are independent of each other and constant in mean and variance over time. Therefore a small p-value is evidence against independence as it would indicate that model residuals are not randomly occurring, but are being influenced by existence of error due to strong bias in the airline delay data. We note that there was no significant autocorrelation of the residuals and the p-values of the Ljung-Box statistic were high for all the airlines, thus implying that all the models derived were adequate for forecasting delay at the airport by airlines.

DISCUSSION

Time series analyses indicate that each airline had a different ARIMA model and it was imminent given that delays for the KLM, SAA, ET, and EA were all stationary;. However, only the British Airways and Kenya Airline had non-stationary delays. In most time series, especially for aviation statistics, such differences or deviations as the non-stationarity in delays is not uncommon. Although we do not attach much importance to this deviation, we note that timeliness performance or delay data for KQ and BA seem to depend on a one lagged time difference. Non-stationarity is just a condition for fitting the Ljung-Box model, by implying that the time series has statistical properties that should not change with time. Appropriate tools such as ACFs and PACFs plots and the Ljung-Box statistics are popular tools for testing if the time series is stationary. In the event that the time series is not stationary, like it was with BA and KQ airline delays, differencing

Table 1. ARIMA models for predicting airline delay at Entebbe International Airport

Airline	Actual Model
KLM	$Y_t = 9.2363$
BA	$Y_t = -1.7171Y_{t-1} -0.8142Y_{t-2} + 0.2849Y_{t-3} + 0.2831Y_{t-4} + 0.1190Y_{t-5} + 0.6804et_{-1} - 0.7919e_{t-21} - 0.8210e_{t-3}$
SAA	Yt = 0.1072Yt-1 + 0.6908Yt-2 - 0.0668Yt-3 - 0.1108Yt-4 + 0.1464Yt-5 - 0.0728et-1 - 0.5910et-2.
ET	$Y_t = 0.0511e_{t-1} + 0.0224e_{t-2} - 0.0139e_{t-3} +0.0493e_{t-4} - 0.0027e_{t-5} + 0.0239e_{t-6} + 0.0626e_{t-7}.$
KQ	$Y_t = -0.2034Y_{t-1} - 0.8119Y_{t-2} + 0.1617Y_{t-3} - 0.0526Y_{t-4} - 0.4224e_{t-1} + 0.5867e_{t-2} - 0.0715e_{t-3}.$
EA	$Y_t = 0.2448Y_{t-1} - 0.0047Y_{t-2} + 0.1277Y_{t-3} – 1.1946e_{t-1} + 0.2016e_{t-2}.$

Note: Y_t refers to current values for delay; $Y_{t-1},\ldots,Y_{t-5}$ refer to delay at periods t-1,…, t-5; e refer to random errors due to moving averages; $e_{t-1},\ldots,e_{t-5}$ refer to random errors at periods t-1,…, t-5 is one of the methods performed on the time series to achieve stationarity in the data. The two airline delays were thus, differenced at least once to make their time series delay data stationary.

The ARIMA models for each of the five airlines were derived and their details presented in Table 1. These models can be used to make airline delay predictions and since the airport delay is a function of airline administered aircraft delay, the models form a credible source of delay information for the airport (Wesonga, 2015).

In terms of autoregressive prediction power, the airline time series with more remote predictive power or influence on the current airline delays shows greater levels of system stability and sustainability compared to those with more recent degrees of prediction. One may attempt to rank the airline efficiency based on the autoregressive degree that shows the British Airways, South African Airways, Kenya Airways, Ethiopian Airways Eagle Air and lastly KLM (Koninklijke Luchtvaart Maatschappij). Other dynamics will be considered to appropriately model and explain the airline delay characteristics of these airlines given that their management is exogenous to the airport under study.

RECOMMENDATIONS

Differences in the time series prediction models of airline delay provide evidence of variations of operational efficiencies between airlines. These differences give a deeper understanding of the connotations of combining airlines in studies that attempt to provide information about aircraft delays at an airport. Understanding these differences is an important ingredient in developing practical solutions that may abate airport based departure and arrival delays (Choo, 2014; Cui, Kuang, Wu, & Li, 2013; Wu & Mengersen, 2013). Thus, the study recommends that studies should attempt to disaggregate delays by airlines due to the unique characteristics they present. By so doing, more reliable solutions will be sought in the face of airline-based delay predictions and operational efficiencies of airports. In all attempts aimed at abating aircraft delays, efforts should be concentrated on improving performance within airlines and subsequently at the airport (Adler & Liebert, 2014; Garrow, Hotle, & Mumbower, 2012; Wesonga, Nabugoomu, & Masimbi, 2013). Further studies are recommended on evaluating time series models on aviation related health outcomes on passengers at the international airports in developing economies.

COMPETING INTEREST

The authors declare that there is no competing interest.

AUTHORS' CONTRIBUTION

The authors equally contributed to the conceptualisation of the study and are responsible for the all technical matters, scientific issues, values, and the manuscript preparation. RW and BM were mainly responsible for data collection, management, analysis drafting of the manuscript and proofreading. FN mainly participated in the conceptualisation, analysis, and proofreading of the final manuscript.

ACKNOWLEDGMENT

We acknowledge the Management of Civil Aviation Authority for the data. We also thank the administration of School of Statistics and Planning for providing a research enabling environment.

REFERENCES

Adler, N., & Liebert, V. (2014). Joint impact of competition, ownership form and economic regulation on airport performance and pricing. *Transportation Research Part A, Policy and Practice*, *64*(0), 92–109. doi:10.1016/j.tra.2014.03.008

Bilotkach, V., & Lakew, P. A. (2014). On sources of market power in the airline industry: Panel data evidence from the US airports. *Transportation Research Part A, Policy and Practice*, *59*(0), 288–305. doi:10.1016/j.tra.2013.11.011

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2013). *Time series analysis: forecasting and control*. John Wiley & Sons.

Choo, Y. Y. (2014). Factors affecting aeronautical charges at major US airports. *Transportation Research Part A, Policy and Practice*, *62*(0), 54–62. doi:10.1016/j.tra.2014.02.006

Civil Aviation Authority. (2012). *Civil Aviation Authority, Uganda - Annual Statistics*. Kampala: Civil Aviation Authority, Ugandao.

Cui, Q., Kuang, H.-B., Wu, C.-Y., & Li, Y. (2013). Dynamic formation mechanism of airport competitiveness: The case of China. *Transportation Research Part A, Policy and Practice*, *47*(0), 10–18. doi:10.1016/j.tra.2012.10.021

Gano, C., & Banavar, S. (2005). National Airspace System Delay Estimation Using Weather Weighted Traffic Counts. In A. I. A. A. Guidance (Ed.), *Navigation, and Control Conference and Exhibit*. American Institute of Aeronautics and Astronautics.

Garrow, L. A., Hotle, S., & Mumbower, S. (2012). Assessment of product debundling trends in the US airline industry: Customer service and public policy implications. *Transportation Research Part A, Policy and Practice*, *46*(2), 255–268. doi:10.1016/j.tra.2011.09.009

International Civil Aviation Organization. (1995). *Economics of satellite-based air navigation services: Guidelines for cost/benefit analysis of communications, navigation and surveillance/air traffic management (CNS/ATM) systems*. Montreal: International Civil Aviation Organization.

Kanafani, A., & Ghobrial, A. A. (1985). Airline hubbing: Some implications for airport economics. *Transportation Research Part A, General*, *19*(1), 15–27. doi:10.1016/0191-2607(85)90003-2

Konishi, S., & Kitagawa, G. (2007). *Information Criteria and Statistical Modeling*. New York: Springer-Verlag.

Lam, W. K., Tang, Y. F., Chan, K. S., & Tam, M.-L. (2006). Short-Term Hourly Traffic Forecasts using Hong Kong Annual Traffic Census. *Transportation*, *33*(3), 291–310. doi:10.1007/s11116-005-0327-8

Nihan, N., & Holmesland, K. (1980). Use of the box and Jenkins time series technique in traffic forecasting. *Transportation*, *9*(2), 125–143. doi:10.1007/BF00167127

Ninj, X. (2007). *Methods for deriving multi factor models for predicting Airport delays*. Fairfax, VA: George Maison University.

Noland, R., Quddus, M., & Ochieng, W. (2008). The effect of the London congestion charge on road casualties: An intervention analysis. *Transportation*, *35*(1), 73–91. doi:10.1007/s11116-007-9133-9

Ryerson, M. S., Hansen, M., & Bonn, J. (2014). Time to burn: Flight delay, terminal efficiency, and fuel consumption in the National Airspace System. *Transportation Research Part A, Policy and Practice*, *69*(0), 286–298. doi:10.1016/j.tra.2014.08.024

Sridhar, B., Grabbe, S. R., & Mukherjee, A. (2008). Modeling and optimization in traffic flow management. *Proceedings of the IEEE*, *96*(12), 2060–2080. doi:10.1109/JPROC.2008.2006141

Voltes-Dorta, A., & Lei, Z. (2014). The impact of airline differentiation on marginal cost pricing at UK airports. *Transportation Research Part A, Policy and Practice*, *55*(0), 72–88.

Wesonga, R. (2015). Airport utility stochastic optimization models for air traffic flow management. *European Journal of Operational Research*, *242*(3), 999–1007. doi:10.1016/j.ejor.2014.10.042

Wesonga, R., & Nabugoomu, F. (2014). Bayesian Model Averaging: An Application to the Determinants of Airport Departure Delay in Uganda. *American Journal of Theoretical and Applied Statistics*, *3*(1), 1–5. doi:10.11648/j.ajtas.20140301.11

Wesonga, R., Nabugoomu, F., & Jehopio, P. (2012). Parameterized framework for the analysis of probabilities of aircraft delay at an airport. *Journal of Air Transport Management*, *23*(0), 1–4. doi:10.1016/j.jairtraman.2012.02.001

Wesonga, R., Nabugoomu, F., & Masimbi, B. (2013). Assessing Aircraft Timeliness Variations By Major Airlines: Passenger Travel Practice In Uganda. [IJSBAR]. *International Journal of Sciences: Basic and Applied Research*, *11*(1), 75–83.

Wu, P. P.-Y., & Mengersen, K. (2013). A review of models and model usage scenarios for an airport complex system. *Transportation Research Part A, Policy and Practice*, *47*(0), 124–140. doi:10.1016/j.tra.2012.10.015

Zhang, Q., Yang, H., Wang, Q., & Zhang, A. (2014). Market power and its determinants in the Chinese airline industry. *Transportation Research Part A, Policy and Practice*, *64*(0), 1–13. doi:10.1016/j.tra.2014.03.003

Zou, B., & Hansen, M. (2014). Flight delays, capacity investment and social welfare under air transport supply-demand equilibrium. *Transportation Research Part A, Policy and Practice*, *46*(6), 965–980. doi:10.1016/j.tra.2012.02.015

Ronald Wesonga is Lecturer of Statistics and Computing at the School of Statistics and Planning, Makerere University. His area of research is modelling, analysis, and development of statistical and computational approaches for complex multidimensional data types in all areas of applications including aviation, health, food security and nutrition, agriculture, weather, and climate. In particular, he has published widely in the area of air traffic management, both in Elsevier and Springer Journals. He is also involved in students' research supervision and teaches Computational Statistics and Applied Statistics at Undergraduate and Graduate levels. He has performed systems analysis and designed a number of systems both online Databases and Electronic Resource Centers for different institutions. Dr. Wesonga obtained his PhD in Statistics, Master of Statistics, and Bachelor of Science in Statistics from Makerere University, Uganda. His focus is to institutionalize functional systems based on statistical and computational evidence necessary for sustainable development for all humanity.

Fabian Nabugoomu is the Deputy Vice Chancellor of Kyambogo University responsible for Finance and Administration (since September 2012). He is an Associate Professor of Statistics and Computing. His area of research engagement is innovative computational approaches for the analysis of complex data, typically multivariate, longitudinal and multi-centred. He also researches in application of ICT in education and management and teaches Applied Mathematics, Computational Statistics and Applied Statistics at Undergraduate and Graduate level. Formerly he was Dean of the Faculty of Science and Technology at Uganda Christian University from 2006 to 2012 where he led and managed the establishment of degree programmes in computer science & IT, health sciences, agricultural sciences, environmental sciences and computational science. He was Senior Lecturer and Head of Mathematics Department at Makerere University from 1998 to 2004. Dr. Nabugoomu obtained his PhD in Statistics from the University of Edinburgh in UK, a Master of Science in Statistics from University of Guelph in Canada and a Bachelor of Statistics degree from Makerere University, Uganda.

Brian Masimbi is a fresh graduate of Master of Statistics. His research interests include modeling airline delay in aviation statistics. He holds Master of Statistics and Bachelor of Statistics of Makerere University.