AIR TRAFFIC FORECAST IN POST-LIBERALIZATION CONTEXT: A DYNAMIC LINEAR MODELS APPROACH

Yesid RODRIGUEZ 1, Wilmer PINEDA 1, Oscar DIAZ OLARIAGA 2*

1Facultad de Estadística, Universidad Santo Tomás, Bogotá, Colombia
2Facultad de Ingeniería Civil, Universidad Santo Tomás, Bogotá, Colombia

Received 22 May 2019; accepted 03 March 2020

Abstract. The process of air transport liberalization in Colombia began in 1991. Liberalization entailed the entry of private capital into the airport sector which subsequently led, in several temporary phases, to the privatization of the country’s main airports. Simultaneously, new air operators entered the market. This new market situation, supported by the complete deregulation of airfares, generated a dynamic and sustained growth of air transport in Colombia for two decades. Within the context of post-liberalization, this article presents a forecast (medium-term – 5 years period) of air traffic in the country’s main airport using DLMs (Dynamic Linear Models). It has the following advantages vs. the usual forecast calculation methodologies: it detects stochastic tendencies that are hidden in the time series. It also detects structural changes that allow estimating the variable effect of exogenous shocks over time without increasing the number of parameters. From the results obtained, it should be noted that the application of DLMs presents MAPE (Mean Absolute Percentage Error) values below 1%, which guarantees predictions of higher accuracy and thus introduces a new alternative model to develop reliable forecasts in air transport, at least in the medium-term.

Keywords: air traffic forecast, liberalization, Dynamic Linear Models, airport, air transport, Colombia.

Introduction

Air transport in Colombia has been developing at an accelerated pace for about two and a half decades. This period coincides with the beginning of a continued implementation of public policies, designed specifically for the sector of air transport to drive and promote it. Since the 1990s, the growth of air traffic in Colombia has been strengthened by the public policy of liberalization of airspace in both domestic and international markets. In addition it has been reinforced by the reorientation of public and private investment toward modernizing and updating airport infrastructure through concessioning the busiest airports in the country (Díaz Olariaga & Ávila, 2015). The first generation of airport concessions was implemented in the mid-1990s and since then three additional generations have taken place (Díaz Olariaga, 2017). In the commercial aviation sector, the national airline was privatized during the same period and new (private) air carriers entered the market, including low-cost carriers (LCC). Since 2012, airfares are completely deregulated (Díaz Olariaga & Zea, 2018).

As a result of public policies, both of privatization and public investment in airport infrastructure (together with deregulating policies in the commercial aviation sector), passenger transport in Colombia has increased 863% (Aerocivil, 2019) during the last 15 years. This significant growth has been boosted and led by the main airport in Colombia, Bogotá–El Dorado International Airport (hereafter BOG), in the country’s capital city. However, the master plans for BOG (as well as several questionable studies and reports) predict growth in demand. Which implies that the airport’s present capacity (as it has no plans to expand) will not meet the anticipated demand. This situation prompted the public sector to approve the construction of a new airport in the outskirts of the city, which is supposed to start operating in 2025/2026.

Forecasts are a crucial aspect of airport planning for determining future capacity requirements. Since airport infrastructure projects are expensive and involve many resources, an understanding based on data for future demand provides planners with necessary information for prospective decision-making in the short, medium and long term. Said data includes the expected number of
aircraft movements, passenger traffic, and air cargo volumes. Therefore, despite unforeseen circumstances, the airport and aviation industry (airlines, aircraft manufacturers and aviation engines, air navigation service providers, etc.) require forecasts to anticipate future scenarios (Kazda & Caves, 2015; de Neufville & Odoni, 2013; Horonjeff et al., 2010). In regards to forecast horizons in the airport industry, it is common to make short-term forecast (next season, within the same year), short-term (1 year), medium-term (5 years), and long-term (20–25 years) (ACI, 2016; ICAO, 2006).

Therefore, the goal of this article is to carry out a forecast for BOG (passengers, air cargo, and air operations or aircraft movements) in the medium-term. To this end, and as a calculation methodology (unique for this type of analysis of air traffic), (DLM) will be used, which in comparison with usual methods for forecast calculation, present the following advantages: detecting stochastic trends hidden in time series (West & Harrison, 2006), as well as structural changes that allow estimating the variable effect in time of exogenous shocks without increasing the number of parameters (Honjo, Shiraki, & Ashina, 2018). Furthermore, the structure of conditional independence (on which the state dynamics are based) allow for an interpretation of forecasts through a recursive algorithm (Petris, Petrone, & Campagnoli, 2009).

The present investigation is organized as follows: in the section Literature Review the current review of existing literature is carried out in two aspects. The first is on the investigations (type and approach) carried out in a context of liberalization of the aviation or air transport industry (worldwide), as it is the framework for the basis of this research. The second aspect focuses on the presentation of research and the methods used by academics to carry out air traffic forecasts. Subsequently the Methodology and Data section describes the methodology used in the research: (DLMs), and the type and origin of the data used in the calculations. In the following section, the Application Case (or Case Study) is presented, i.e. the information and general data of the airport for which the traffic prognosis is developed. In the next and penultimate section, all the results are presented and analyzed. In the concluding section the final results of the investigation are revealed.

1. Literature review

There are many studies that address the various aspects of the liberalization of the air transport market and industry. These studies deal with issues such as the spatial effects of deregulation on connectivity and accessibility (Díaz Olariaga & Carvajal, 2016; O’Connor, 2003; Bowen, 2002, 2000; Ivy, 1995; Chou, 1993), market competition and consolidation (Goetz, 2002; Goetz & Sutton, 1998; Debbage, 1993; Graham, 1993), network structures (Díaz Olariaga & Zea, 2018; O’Kelly, 1998; Shaw and Ivy, 1994; Shaw, 1993), airfare prices (Vowles, 2006; Stavins, 2001; Oum et al., 1996), and airline alliances (Fan et al., 2001; Oum, Yu, & Zhang, 2001; Vowles, 2000; Youssef & Hansen, 1994). Other studies focus on the analysis of diverse situations (e.g. demand behavior), in the post liberalization context, particularized in certain countries or regions; in addition to the countless jobs in relation to the United States, we have, for example, from Europe (Eriksen & Pettersson, 2012; Fan, 2006; Goetz & Graham, 2004; Thompson, 2002; Graham, 1998, 1997; Ivy, 1995; Dennis, 1994), from Canada (Oum, 1991), from Australia (Hooper, 1998; Forsyth, 1991), from Asia (Bowen, 2000; Bowen & Leinbach, 1995), from Brazil (Oliveira, Lohmann, & Costa, 2016; Koo & Lohmann, 2013), from Central and Eastern Europe (Jankiewicz & Huderek-Glapska, 2015), from Africa (Njoya, Christidis, & Nkita, 2018; Abate, 2016; Dobruszkes, Mondou, & Ghedira, 2016; Njoya, 2015; Surovitskikh & Lubbe, 2015; Ismaila, Warnock-Smith, & Hubbard, 2014; Daramola & Jaja, 2011), and finally from Colombia, where there is only one related research, that of Díaz Olariaga et al. (2017). From the perspective of public policy, Koo and Lohmann (2013) examine the relationship between the volatility of aeronautical public policy and the spatial evolution of the air transport supply. These authors focus on the domestic aviation sector of two comparative cases: Australia and Brazil. On the other hand, Rolim et al. (2016) analyze the development of demand in recently privatized airports, such as in Brazil. Changes in traffic concentration at airports as a result of liberalization have also been analyzed (Díaz Olariaga & Zea, 2018; Koo, Tan, & Duval, 2013; Halpern, 2011; Derudder & Witlox, 2009; Dobruszkes, 2009; Papatheodorou & Arvanitis, 2009).

Regarding the methods for forecasting, the air transport industry has been addressing the issue of traffic prognosis for at least six decades. However, academics only began to present formal studies and research about three decades ago. During this time a variety of models have been developed to predict the demand of passengers. The most used prediction methods can be classified into two large groups: economic models and time-series models (Dantas, Oliveira, & Repolho, 2017). The economic methods focus on the correlation between the demand of passengers and multiple variables, which are considered to be influential in the change of the economic environment and traffic system. The forecast models are established through a series of equations. Commonly used models include regression analysis (Abed, Ba-Fail, & Jasimuddin, 2001), causality test (Fernandes & Pacheco, 2010), logit model (Garrow & Koppelman, 2004), and gravitational models (Grosche, Rothlauf, & Heinzl, 2007). Time series methods primarily rely on historical data to predict by extracting the intrinsic relationship between current data and past observations. The various time-series models have been used to forecast passenger demand, such as smoothing techniques (Samagiao & Wolters, 2010), the adapted Markov model (Chin & Tay, 2001), ARIMA/SARIMA (Tsui et al., 2014), seasonal adjustment method
(Aston & Koopman, 2006), etc. However, due to the non-linear nature of passenger demand, economic and time-series approaches are severely criticized for their limited and ineffective forecasting (Tsui et al., 2014). Therefore, some academics try to explore other methods, such as artificial intelligence (for example, neural networks), which is characterized by self-adaptation and non-linearity and can map arbitrary functions (Jin et al., 2020; Xiao et al., 2014). Regarding the methodology used in the present investigation DLMs virtually no related publications have been found. Thus the interest and motivation to test this technique to evaluate behavior and reliability in the prediction of typical air traffic variables and their positive benefits (cited in the Introduction and more developed in the Methodology section).

2. Methodology and data

In any statistical application, a crucial and challenging step is to carefully specify the model. The first strategy is a static model, where the effect of time does not play a prominent role. For this research, Dynamic Models (DMs) have been chosen. Unlike static models, some elements that participate in the construction of the model do not remain invariant but are considered as functions of time describing temporal trajectories (Glynn et al., 2019; Laine, 2019; McAlinn & West, 2019; Pole, West, & Harrison, 2018).

DMs have the advantage of having "dynamics" in the model’s parameters, thereby rendering the parameters not fixed but changing or dependent on time. Their main application is the analysis of time series. They also have the advantage of being useful to perform sequential analyses because the updating of parameters is carried out based on the data that has been obtained sequentially.

The development of forecasts is usually based on autoregressive models, moving averages or their combination. However, such models have a complicated verisimilitude function and due to this the final distribution of parameters inherit the same difficulty. Based on the aforementioned, DLMs, which are a particular case of DMs, are used for modeling time series in order to carry out forecasts by distributions of stochastic variables that influence observations in time. One of their advantages is that by using them one realizes that they are simpler models, yet powerful enough to adjust and forecast data and they may include explanatory variables in a simple way (Sargan & Bhargava, 1983; Ahn & Schmidt, 1995; Arellano & Bond, 1991; Arellano & Bover, 1995; Gelman et al., 2013; Kenkel, 2018). DLMs are defined under the following structure for each time t (Valencia & Correa, 2013; Bolstad, 2007; Glynn et al., 2019; Asparouhov, Hamaker, & Muthén, 2018). Let $Y_t = (Y_{1,t}, Y_{2,t}, ..., Y_{k,t})'$ represent observations at time t. Hence each $Y_t$ is a $k \times 1$ vector. Furthermore, let $\theta_t = \left[ \theta_{1,t}, \theta_{2,t}, ..., \theta_{q,t} \right]'$ be the true $q$ processes of interest. We express the DLM with the following equations:

Observation equation: $Y_t = F_t \theta_t + V_t$,  \hspace{1cm} (1)

System equation: $\theta_t = G_t \theta_{t-1} + W_t$,  \hspace{1cm} (2)

where:

$F_t$ is a matrix of a dynamic regression.

$G_t$ is a matrix of a state.

$V_t$ is a matrix of a observational variance, $V_t \sim N(0, V_t)$

$W_t$ is a matrix of evolution, $W_t \sim N(0, W_t)$

$\theta_t$ is a vector of parameters.

In time 0 a priori distribution is postulated for $(\theta_0 | D_0)$ where $D_0$ represents available information until time zero. West and Harrison (2006) suggest $(\theta_0 | D_0) \sim N(\mu_0, C_0)$, where $\mu_0$ and $C_0$ are the vector of averages and the matrix of variances and covariances, respectively.

The observation equation defines the observational model for an answer $Y_t$, and its relation with p covariables or explanatory variables $F_t$. The first explanatory variable is generally a constant or intercept that represents the level of the series. As $F_t$ is univariate, then $\theta_t$ is a vector of the form $(\theta_{0,t}, \theta_{1,t}, ..., \theta_{p-1,t})'$. It is possible to consider $Y_t$ as multivariate, in which case $\theta_t$ is a matrix of dimension $m \times p$. $m_0$ it’s a vector of zeros and $C_0$ it is the diagonal matrix of variances and covariances.

The system equation presents the evolution of the parameters in time. If the model includes p changing coefficients, it will result in the evolution to be defined as a transition matrix $G_t$ of dimension $p \times p$, where $p$ represents the number of covariates in each of the models.

Finally, DLMs present errors $v_t$ and $w_t$, with variances dependent on time $V_t$ and $W_t$ that denote the matrix of observational variance and the evolution of variance, respectively.

When dealing with temporal series, it is important to consider that, for DLMs, a source of variability that works for representing errors in the observation equation and in the system’s equation is known as a vector of permanent effects. Even though it appears to have more limitations, some classical models of time series are presented as a particular case, especially ARMA models, for theoretical relationships of dynamic linear models with ARIMA models see Durbin and Koopman (2012), Tsui et al. (2014), Box et al. (2016) and Wei (2006). They are dealt with through Kalman’s filter, when the error terms in the observation equation follow a normal distribution, they are independent and are distributed identically in average 0 and known variance.

For the proposed models, 5 chains of the MCMC (Markov Chain Monte Carlo) sampler were made, with 100,000 iterations each, discarding the first 10,000 first iterations, using the JAGS (Just Another Gibbs Sampler) methodology (Plummer, 2003). To determine the convergence of the chains, the Gelman-Rubin R-hat statistic (Gelman et al., 2013) will be used using the "coda" library of the R software (Plummer et al., 2006), which to present values below 1.1.
In order to determine the strength in numerical terms of the proposed model, Mean Absolute Percentage Error (MAPE) will be used, which measures the size of error (absolute) in percentage terms. The fact that the magnitude of percentage error is estimated, it renders it an indicator frequently used by forecast developers due to its easy interpretation. A small MAPE value indicates that forecasts have a higher likelihood of being accurate (S. Kim & H. Kim, 2016; Ren & Glasure, 2009).

There is data available about air traffic in the airport of study (passengers, air cargo and operations or air movements) during the last four decades (1979–2017) (Aerocivil, 2019). Likewise, socioeconomic data is available for the city where the airport is located (GDP, GDP/per capita, population, etc.) (DANE, 2019; Banco de la República de Colombia, 2019). According to the chosen variables as covariables, a medium-term forecast will be presented due to the changing economic conditions and their effects on air traffic. For that purpose, years 2018 to 2022 will be forecasted. To achieve such forecast, first forecasts should be carried out using ARIMA models (Brooks, 2008) on the covariables chosen in order to include these new variables in the selected model, thereby attempting to obtain a relatively low MAPE.

3. Application case

In Colombia, the aviation industry has been liberalized since the beginning of the 1990s and airfares are completely deregulated since 2012. Within the national context, the case of Bogotá-El Dorado International Airport (IATA code: BOG; OACI code: SKBO) has been selected. This is the main airport in the country and the main country hub, situated in the city of Bogotá (capital of Colombia, and with more than 8 million inhabitants), about 7.5 miles from the city center. The airport is a public property but it has been concessioned to the private sector since 2007 (Díaz Olariaga, 2017), a year when the airport developed its first significant expansion in infrastructure and facilities (with an investment of $650 million), which was finished in 2013. In 2015, a second expansion began which finished at the end of 2018. About 25,000 people work at the airport. Total traffic data (for 2017): 30M Pax, 690,000 Tn air cargo, 320,000 operations (or aircraft movements) (Aerocivil, 2019).

4. Results

Bayesian statistics present an ideal alternative to make models without the problem of updating data presented by classical statistics. One of the advantages is obtaining the new information, since a posteriori distribution can be updated and will be used as an a priori distribution, obtaining a new, more updated a posteriori distribution. This is a great advantage of Bayesian analysis because classical analysis requires everything to be recalculated as more data appears (Bolstad, 2007). In the calculation of the different forecasts for the auxiliary variables, ARIMA models were used and subsequently the DLMs, with their respective MAPE analysis, to choose the best model for each of the variables.

In the case of the variable “national (or domestic) passengers” Consumer Price Index (CPI) was used as an auxiliary variable to estimate the future forecast. Figure 1 shows the result.

In Figure 1 Model 1 is presented, where the behavior of estimated values for the chosen model is shown. These values overlap with the behavior of the original values. 1.08% MAPE can also be observed (see Table 1). To estimate the forecast, ARIMA (2,1,0) model was used in the variable CPI in order to carry out 5 years forecast and for it to be included in the variable of national passengers.

The final model has the following structure:

\[
y_t = \beta_0 + \beta_1 CPI_t + v_t ;
\]

\[
\beta_0 \sim N(\beta_{0,t-1}, \frac{1}{\hat{h}_1}) ;
\]

\[
\beta_1 \sim N(\beta_{1,t-1}, \frac{1}{\hat{h}_2}) ;
\]

\[
\theta_t = \left( \frac{\beta_{0,t}}{\beta_{1,t}} \right);
\]

\[
F_t = \begin{bmatrix} 1 & CPI_t \end{bmatrix};
\]

\[
G_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]

In the case of the variable “international passengers” GDP, Population and Currency Exchange Rate (in Spanish TRM) were used as auxiliary variables to estimate the future forecast, thereby obtaining the results shown in Model 2 (see Figure 2).

![Figure 1. Model 1, forecast for the variable "domestic passengers"](image)

Table 1. Comparison of MAPE values for Model 1

| Model | MAPE   | Variables                          |
|-------|--------|-----------------------------------|
| 1     | 0.01082| National passengers, CPI           |
| 2     | 0.01630| National passengers, GDP per capita|
| 3     | 0.02560| National passengers, national passengers with delay t-1 |
| 4     | 0.02884| National passengers, per capita GDP, CPI |
In Figure 2 Model 2 is presented, where the behavior of estimated values for the model chosen is shown. These values overlap with the behavior of the original values. 0.97% MAPE can also be observed (see Table 2). To estimate the forecast, ARIMA (3, 1, 0) model was used in the variable GDP, ARIMA (1, 1, 0) model in the variable population, and ARIMA (2, 1, 0) model in the variable TRM in order to carry out 5 years forecast and for it to be included in the variable of international passengers.

The final model has the following structure:

\[
y_t = \beta_0 + \beta_1 GDP_t + \beta_2 population_{t-1} + \beta_3 TRM_{t-1} + v_t;
\]

\[
\beta_0 \sim N(\beta_{0,t-1}, \frac{1}{h_1});
\]

\[
\beta_1 \sim N(\beta_{1,t-1}, \frac{1}{h_2});
\]

\[
\beta_2 \sim N(\beta_{2,t-1}, \frac{1}{h_2});
\]

\[
\beta_3 \sim N(\beta_{3,t-1}, \frac{1}{h_3});
\]

\[
\theta_t = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix};
\]

\[
F_t = \begin{bmatrix} 1 & CPI_{t-1} & population_{t-1} & TRM_{t-1} \end{bmatrix};
\]

\[
G_t = \begin{pmatrix} 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 1 \end{pmatrix}_{4 \times 4}.
\]

In the case of the variable “operations” (take-offs/landings, where national and international operations are included), GDP per capita, Population and Currency Exchange Rate (in Spanish TRM) were used as auxiliary variables to estimate the future forecast, thereby obtaining the results shown in Model 3 (see Figure 3). In Figure 3 Model 3 is presented, where the behavior of estimated values for the chosen model is shown. These values overlap with the behavior of the original values. 0.24% MAPE can also be observed (see Table 3). To estimate the forecast, ARIMA (3, 1, 0) model was used in the variable GDP per capita, ARIMA (1, 1, 0) model in the variable population, and ARIMA (2, 1, 0) model in the variable TRM in order to carry out 5 years forecast and for it to be included in the variable of operations.

The final model has the following structure:

\[
y_t = \beta_0 + \beta_1 GDP_{t-1} + \beta_2 population_{t-1} + \beta_3 TRM_{t-1} + v_t;
\]

\[
\beta_0 \sim N(\beta_{0,t-1}, \frac{1}{h_1});
\]

\[
\beta_1 \sim N(\beta_{1,t-1}, \frac{1}{h_1});
\]

\[
\beta_2 \sim N(\beta_{2,t-1}, \frac{1}{h_2});
\]

\[
\beta_3 \sim N(\beta_{3,t-1}, \frac{1}{h_3});
\]

\[
\theta_t = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix};
\]

\[
F_t = \begin{bmatrix} 1 & CPI_{t-1} & population_{t-1} & TRM_{t-1} \end{bmatrix};
\]

\[
G_t = \begin{pmatrix} 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 1 \end{pmatrix}_{4 \times 4}.
\]

Table 2. Comparison of MAPE values for Model 2

| Model | MAPE   | Variables                                |
|-------|--------|------------------------------------------|
| 1     | 0.0097555 | International passengers, GDP, population, TRM |
| 2     | 0.0111818 | International passengers, TRM             |
| 3     | 0.0114201 | International passengers, population      |
| 4     | 0.0122443 | International passengers, GDP             |
| 5     | 0.0137127 | International passengers, GDP, TRM        |
| 6     | 0.0156288 | International passengers, GDP             |
| 7     | 0.0163126 | International passengers, population      |
| 8     | 0.0195899 | International passengers, international passengers with delay t-1 |

Table 3. Comparison of MAPE values for Model 3

| Model | MAPE   | Variables                                |
|-------|--------|------------------------------------------|
| 1     | 0.00245412 | Operations, GDP per capita, population, TRM |
| 2     | 0.00259602 | Operations, population, TRM              |
| 3     | 0.00266953 | Operations, GDP per capita, TRM          |
| 4     | 0.00418937 | Operations, TRM                         |
| 5     | 0.00439192 | Operations, Operations with delay t-1    |
| 6     | 0.0047898  | Operations, GDP per capita               |
| 7     | 0.00646659 | Operations, GDP per capita, population   |
| 8     | 0.00839123 | Operations, population                   |
In the case of the variable “national (or domestic) air cargo”, GDP per capita and population were used as auxiliary variables to estimate the future forecast, thereby obtaining the results shown in Model 4. In Figure 4 Model 3 is presented, where the behavior of estimated values for the chosen model is shown. These values overlap with the behavior of the original values. 0.42% MAPE can also be observed (see Table 4). To estimate the forecast, ARIMA (3, 1, 0) model was used in the variable GDP per capita GDP and ARIMA (1, 1, 0) model in the variable population in order to carry out 5 years forecast and for it to be included in the variable national air cargo.

The final model has the following structure:

\[ y_t = \beta_0 + \beta_{1i}GDP_{- per \_ capita i} + \beta_{2i}population_{2i} + \nu_t; \]  

\[ \beta_0 \sim N(\beta_{0,i-1}; \frac{1}{h_i}); \]  

\[ \beta_1 \sim N(\beta_{1,i-1}; \frac{1}{h_2}); \]  

\[ \beta_2 \sim N(\beta_{2,i-1}; \frac{1}{h_2}); \]  

\[ \beta_3 \sim N(\beta_{3,i-1}; \frac{1}{h_3}); \]  

The model in the case of the variable “national (or domestic) air cargo”, GDP and international trade (imports and exports) were used as auxiliary variables to estimate the future forecast, thereby obtaining the results shown in Model 5 (see Figure 5).

In Figure 5 Model 5 is presented, where the behavior of estimated values for the chosen model is shown. These values overlap with the behavior of the original values. 0.63% MAPE can also be observed (see Table 5). To estimate the forecast, ARIMA (3, 1, 0) model was used in the variable GDP, ARIMA (1, 1, 0) model in the variable imports, and ARIMA (1, 1, 0) model in the variable exports in order to carry out 5 years forecast and for it to be included in the variable international air cargo.

Table 4. Comparison of MAPE values for Model 4

| Model | MAPE  | Variables                                |
|-------|-------|------------------------------------------|
| 1     | 0.0042948 | National air cargo, GDP per capita , population |
| 2     | 0.00438331 | National air cargo, population          |
| 3     | 0.00951284 | National air cargo, GDP per capita       |
| 4     | 0.01479821 | National air cargo, national air cargo with delay t-1 |

Table 5. Comparison of MAPE values for Model 5

| Model | MAPE      | Variables                                |
|-------|-----------|------------------------------------------|
| 1     | 0.00633563 | International air cargo, imports, exports |
| 2     | 0.00750094 | International air cargo, GDP, exports     |
| 3     | 0.00762745 | International air cargo, GDP, exports     |
| 4     | 0.00840434 | International air cargo, imports, exports |
| 5     | 0.00891068 | International air cargo, international air cargo with delay t-1 |
| 6     | 0.01684195 | International air cargo, exports          |
| 7     | 0.02385259 | International air cargo, imports          |
| 8     | 0.02820351 | International air cargo, GDP              |
The final model has the following structure:

\[ y_t = \beta_0 + \beta_1 \text{imports}_{t-1} + \beta_2 \text{exports}_{2t} + v_t; \]  
\[ \beta_0 \sim N(\beta_{0,t-1}, \frac{1}{R_1}); \]  
\[ \beta_1 \sim N(\beta_{1,t-1}, \frac{1}{R_2}); \]  
\[ \beta_2 \sim N(\beta_{2,t-1}, \frac{1}{R_2}); \]  
\[ \theta_t = \begin{bmatrix} \beta_{0,t-1} \\ \beta_{1,t-1} \\ \beta_{2,t-1} \end{bmatrix}; \]  
\[ F_t = \begin{bmatrix} 1 & \text{imports}_{t-1} & \text{exports}_{2t} \end{bmatrix}; \]  
\[ G_t = \begin{bmatrix} 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 1 \end{bmatrix}_{3 \times 3}. \]  

Conclusions

Air traffic forecasts are essential in airport planning for determining future capacity requirements. Since airport infrastructure projects are expensive and involve many resources, an understanding based on data from future demand provides airport planners with the necessary information for effective decision-making in the short, medium, and long term. Therefore, regardless of unforeseen circumstances, the airport industry requires forecasts to anticipate future scenarios.

Considering the advantages of using DLMs in the forecast of time series, an initial description of variables was made, revealing a growing behavior as well as strong correlations in time with the covariables. Regarding the covariables presented in the models, an ARIMA model was used to carry out the future forecast and the values to be included in the model chosen. In order to test the convergences of the chains, the R-hat test was applied, which showed values in the test of R-hat < 1.1 in all the final models chosen. The result of the application of DLMs presents MAPE values below 1%, which ensures high predictability forecasts. Furthermore, it could be verified that when the model chosen is contrasted with models that compared the variable with delay t-1 (which is equivalent to AR(1) models), DLMs showed an acceptable performance as alternative models to develop reliable forecasts in air transport (or air traffic prognosis), at least in the medium term.

As seen in the results obtained, the DLMs showed excellent performance by giving a new, and to some extent, original alternative to develop reliable forecasts in air traffic (no investigations were found that use the DLMs for calculating forecasts in the field of air traffic). However, the present investigation focused on a medium-term prognosis (5 years), and with an important historical series (39 years), thus the next phase of research will be to test the performance and reliability of the method for calculating a forecast in a long-term period (20–25 years – usual requirement of airport planners) by using a similar historical series.

References

Abate, M.A. (2016). Economic effects of air transport market liberalization in Africa. Transportation Research Part A, 92, 326–337. https://doi.org/10.1016/j.tra.2016.06.014

Abed, S. Y., Ba-Fail, A. O., & Jasmuddin, S. M. (2001). An econometric analysis of international air travel demand in Saudi Arabia. Journal of Air Transport Management, 7, 143–148. https://doi.org/10.1016/S0969-6997(00)00043-0

ACI. (2016). Guide to World Airport Traffic Forecasts. Montreal: Airports Council International.

Aerocivil. (2019). Statistics. http://www.aerocivil.gov.co/estadisticas.html

Ahn, S., & Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. Journal of Econometrics, 68, 5–27. https://doi.org/10.1016/0304-4076(94)01641-C

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies, 58(2), 277–297. https://doi.org/10.2307/2297968

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of Error-Components Models. Journal of Econometrics, 68, 29–51. https://doi.org/10.1016/0304-4076(94)01642-D

Asparouhov, T., Hamaker, E., & Muthén, B. (2018). Dynamic structural equation models. Structural Equation Modeling: A Multidisciplinary Journal, 25(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Aston, J. A. D. & Koopman, S. J. (2006). A non-Gaussian generalization of the airline model for robust seasonal adjustment. Journal of Forecasting, 25, 325–349. https://doi.org/10.1002/for.991

Banco de la República de Colombia. (2019). Estadísticas. http://www.banrep.gov.co/es-es/estadisticas

Bolstad, W. (2007). Introduction to Bayesian statistics. Wiley. https://doi.org/10.1002/9780470181188

Box, G., Jenkins, G., Reinsel, G., & Ljung, G. (2016). Time series analysis: forecasting and control. John Wiley & Sons.

Bowen, J. (2002). Network change, deregulation, and access in the global airline industry. Economic Geography, 78(4), 425–439. https://doi.org/10.1017/S0884817302000643

Bowen, J. (2000). Airline hubs in Southeast Asia: national economic development and nodal accessibility. Journal of Transport Geography, 8(1), 25–41. https://doi.org/10.1006/jotr.2000.0030-7

Bowen, J., & Leinbach, T. (1995). The state and liberalization: the evolution of American Geographers. American Journal of Sociology, 85(3), 468–493. https://doi.org/10.1111/j.1467-8306.1995.tb01809.x

Brooks, C. (2008). Introductory econometrics for finance. Cambridge (UK): Cambridge University Press. https://doi.org/10.1017/CBO9780511841644

Chin, A. T. H., & Tay, J. H. (2001). Developments in air transport market liberalization. Transportation Research Part A, 35(7), 319–330. https://doi.org/10.1016/S0969-6997(00)00026-6

Chou, Y. H. (1993). Airline deregulation and nodal accessibility. Journal of Transport Geography, 1(1), 36–46. https://doi.org/10.1016/0966-6923(93)90036-Y
Valencia, M., & Correa, J. (2013). Un Modelo Dinámico Bayesi-
nano para el Pronóstico de Energía Diaria. Revista Ingeniería
Industrial, 12(2), 7–17.
Vowles, T. (2006). Airfare pricing determinants in hub-to-hub
markets. Journal of Transport Geography, 14(1), 15–22.
https://doi.org/10.1016/j.jtrangeo.2004.10.004
Vowles, T. (2000). The geographic effects of US airline alliances.
Journal of Transport Geography, 8(4), 277–285.
https://doi.org/10.1016/S0966-6923(00)00012-0
Wei, W. (2006). Time series analysis univariate and multivariate
methods. Pearson Addison Wesley.
West, M., & Harrison, J. (2006). Bayesian forecasting and dynam-
ic models. Springer Science & Business Media.
Xiao, Y., Liu, J. J., Hu, Y., Wang, Y., Lai, K. K., & Wang, S. (2014).
A neuro-fuzzy combination model based on singular spec-
trum analysis for air transport demand forecasting. Journal
of Air Transport Management, 39, 1–11.
https://doi.org/10.1016/j.jairtraman.2014.03.004
Youssef, W., & Hansen, M. (1994). Consequences of strategic al-
liances between international airlines: the case of Swissair and
SAS. Transportation Research A, 28(5), 415–431.
https://doi.org/10.1016/0965-8564(94)90024-8