

Air transport forecast in post-liberalization context: A Dynamic Linear Models approach



- ▶ Introduction
- ▶ Literature review
- ▶ Methodology and data
- ▶ Application case
- ▶ Results
- ▶ Conclusions
- ▶ References





- ▶ Air transport in Colombia has been developing at an accelerating and dynamic 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. The growth of air traffic in Colombia has been strengthened since the 1990s by the public policy of liberalization of airspace in both domestic and international markets and by the re-orientation of public and private investment toward modernizing and updating airport infrastructure through concessioning the busiest airports in the country [1].





- ▶ Therefore, the goal of this article is to carry out a forecast for BOG (passengers, air cargo, and air operations or movements) in the short term. To this end, and as a calculation methodology (novel for this type of analysis of air traffic), Dynamic Linear Models (DLM) will be used, which, in comparison with usual methods for forecast calculation, presents the following advantages: it detects stochastic trends hidden in time series [5], as well as structural changes that allow estimating the variable effect in time of exogenous shocks without increasing the number of parameters [6].





- ▶ There is much research that deals with different aspects of liberalization of aviation / air transport. Many studies address such topics as spatial effects in deregulating connectivity and accessibility [8], market competition and consolidation [9], configuration network structures ([3] [10]), price of air tickets [11], and airline alliances [12]. Other studies focus their attention on the analysis of different situations (e.g. the behavior of demand) within the post-liberalization context in some given countries or regions [13].





- ▶ On the other hand, Rolim et al. [14] analyze the development of demand in recently privatized airports, as in the case of Brazil. The changes in traffic concentration in airports as a result of liberalization have also been discussed [15]. For instance, Sun and Schonfeld [16] use stochastic programs to estimate, on the one hand, traffic future demand, and, on the other hand, how to optimize decision making about the future development of an airport (investments in capacity), where uncertainty in traffic forecast is considered.





- ▶ All this is applied to markets or countries where air transport is completely liberalized [17], by using systems dynamics they carry out a forecast of transport demand of air cargo to determine the capacity of the cargo terminal and its expansion (the study included a set of airports in Taiwan). Singh et al. [18], through an econometric model, conduct a forecast study of both air traffic and investment in airport capacity for a 20 years period for the airport system in India and within the context of post-liberalization in the air traffic industry.





- ▶ In any statistical application, a crucial and sometimes difficult step is to carefully specify the model. The first strategy is a static model, where the effect of time does not play an important role. For this research, Dynamic Models (DMs) have been chosen because, unlike the case of static models, some of the elements that participate in the construction of the model do not remain invariable, but are considered as functions of time, describing temporal trajectories [22].





- ▶ Dynamic Models (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 have been obtained sequentially.





The development of forecasts is usually based on models of the autoregressive type, moving averages or their combination. However, such models have a complicated verisimilitude function and, therefore, the final distribution of parameters inherit the same difficulty. Based on the aforementioned, Dynamic Linear Models (DLMs), which are a particular case of Dynamic Models (DMs), are used for modeling time series in order to carry out forecasts by distributions of stochastic variables that influence observations in time





There are data available about the air traffic in the airport under study (passengers, air cargo and operations or air movements) during the last four decades (1979-2017) [34]. Likewise, socioeconomic data are available as regards the city where the airport is located (GDP, GDP/per capita, population, etc.) [35] [36]. For that purpose, the years 2018 to 2022 will be forecast. To achieve such forecast, first forecasts should be carried out using ARIMA models [37] on the covariables chosen in order to include these new variables in the selected model, thereby attempting to obtain a relatively low MAPE.





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 are accurate and that they will have a higher likelihood of being accurate forecasts [32] [33].





- ▶ DLMs are defined under the following structure for each time t [29] [30] [22] [31]:

Observation equation: $\mathbf{Y}_t = \mathbf{F}_t' \boldsymbol{\theta}_t + \mathbf{v}_t, \mathbf{v}_t \sim N(\mathbf{0}, \mathbf{V}_t)$

System equation: $\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \mathbf{w}_t \sim N(\mathbf{0}, \mathbf{W}_t)$

where:

\mathbf{F}_t is a matrix of a known dynamic regression.

\mathbf{G}_t is a matrix of a known state.

\mathbf{V}_t is a matrix of a known observational variance.

\mathbf{W}_t is a matrix of known evolution.

$\boldsymbol{\theta}_t$ is a vector of parameters.





- ▶ 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) is chosen. This is the main airport in the country and 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.



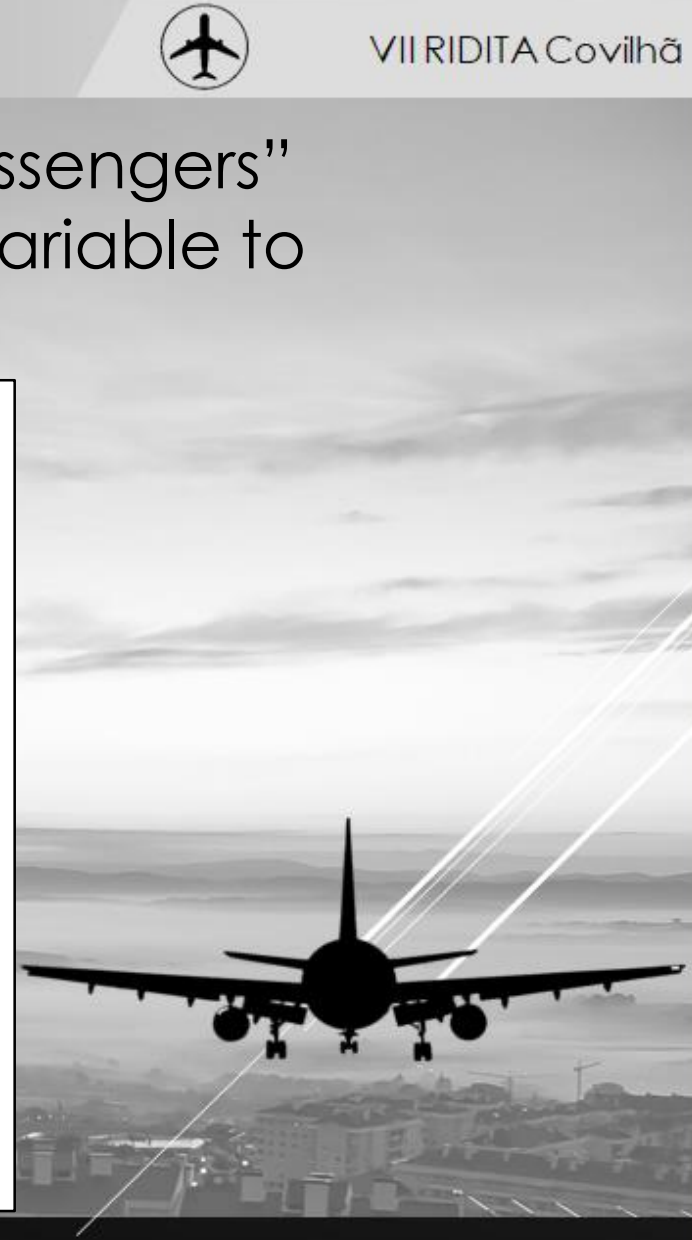
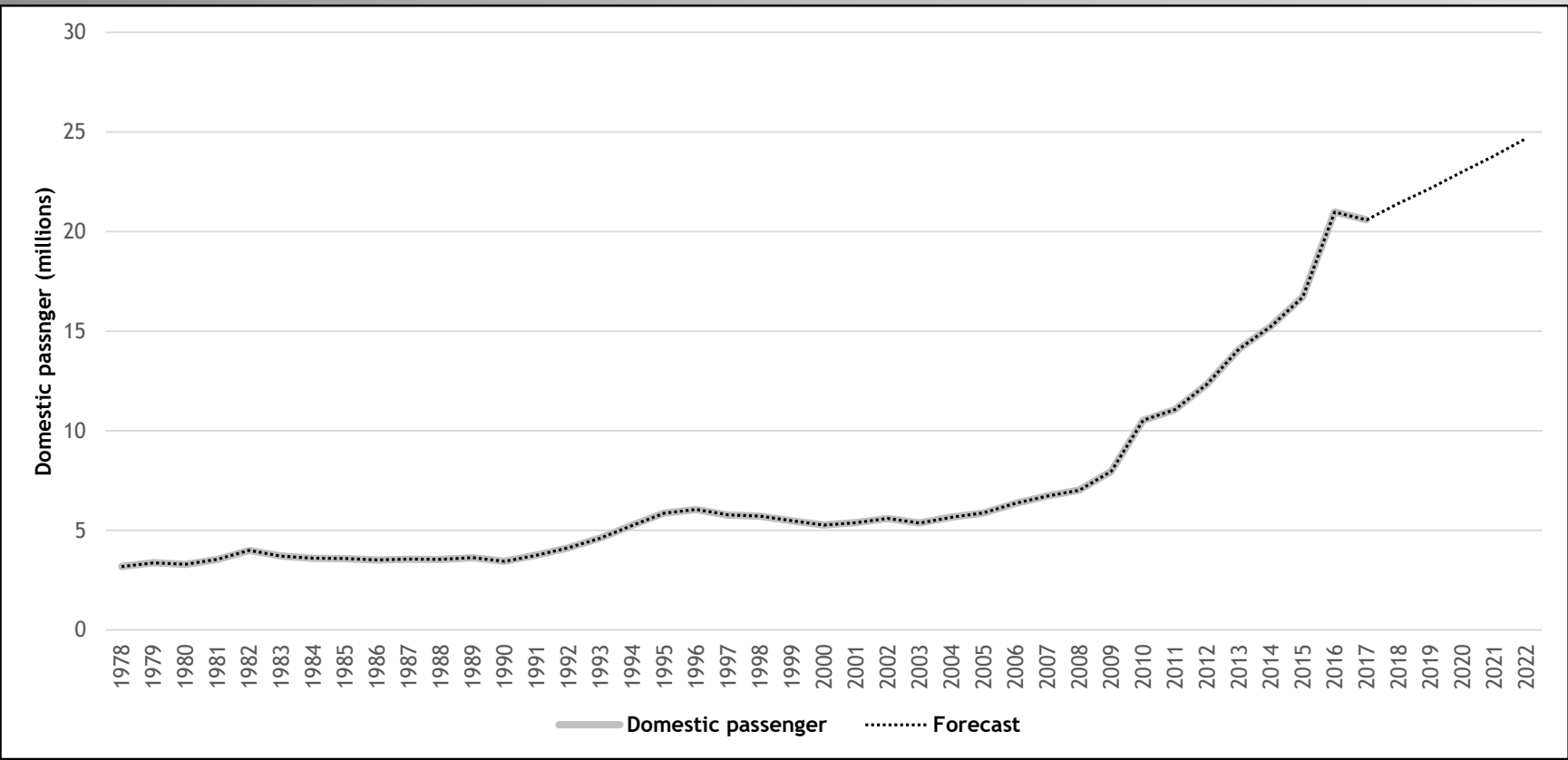


- ▶ The airport is public property but it has been concessioned to the private sector since 2007 [2], a year when the airport developed a first (and significant) expansion in infrastructure and facilities (with an investment of USD 650 million), which finished in 2013. In 2015 a second expansion began, which finished at the end of 2018. About 25,000 people work at the airport. BOG is the third terminal for passenger transport and the first for air cargo in Latin America [1].





► 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 Fig. 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.

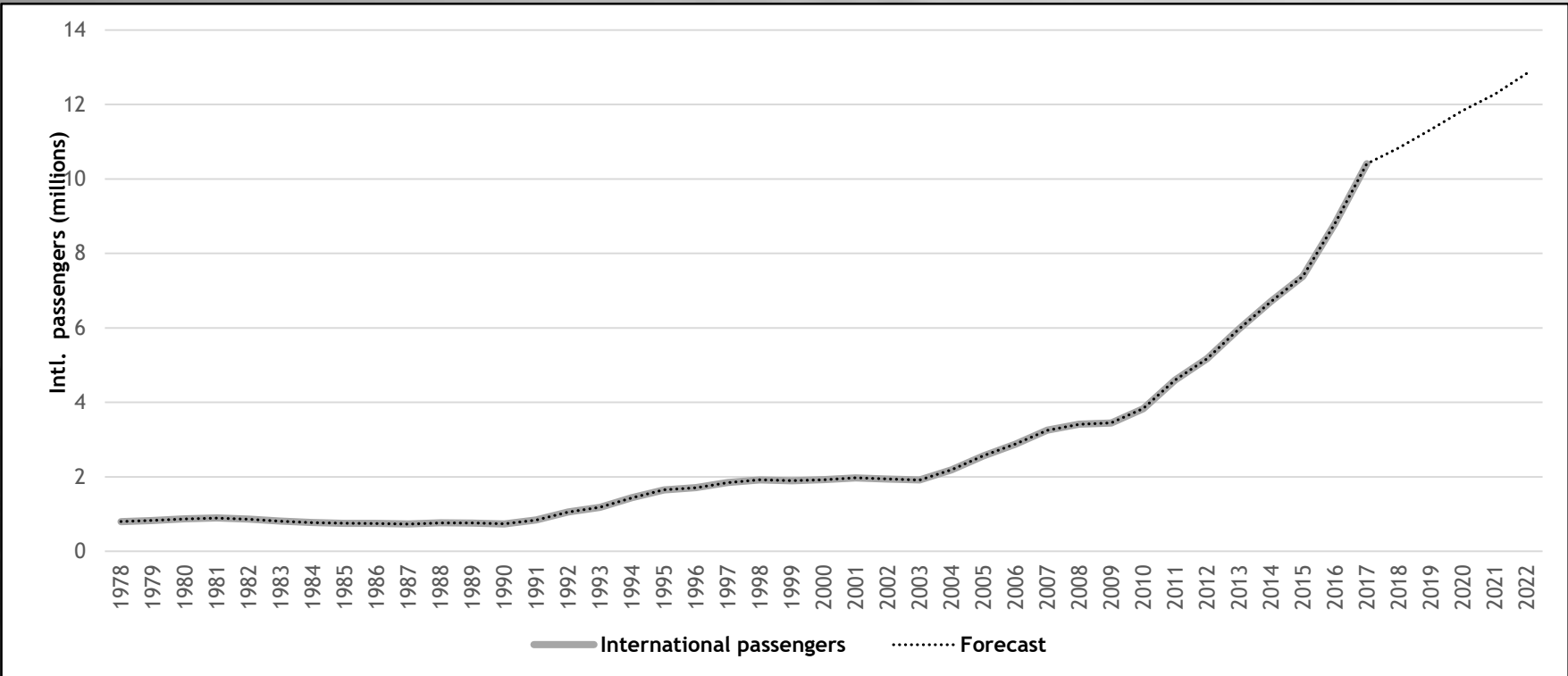
Table 1. Comparison of MAPE values for Model 1. Source: Authors.

Modelo	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 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 Fig. 2).





- ▶ In Fig. 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.

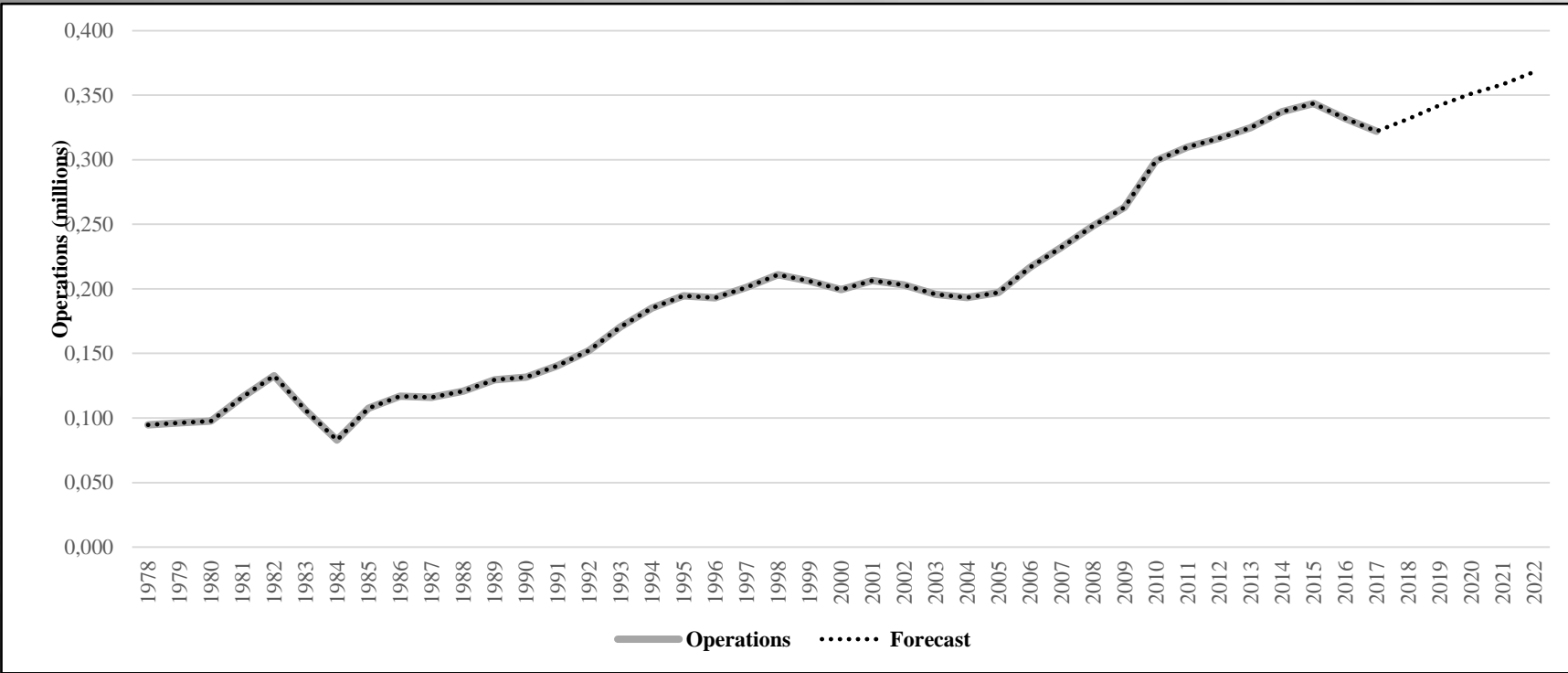
Table 2. Comparison of MAPE values for Model 2. Source: Authors.

Model	MAPE	Variables
1	0,0097555	International passengers, GDP, population, TRM
2	0,0111818	International pasengers, 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, population
7	0,0163126	International passengers, GDP
8	0,0195899	International passengers, international passengers with delay t-1





► 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 Fig. 3).



- ▶ In Fig. 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.

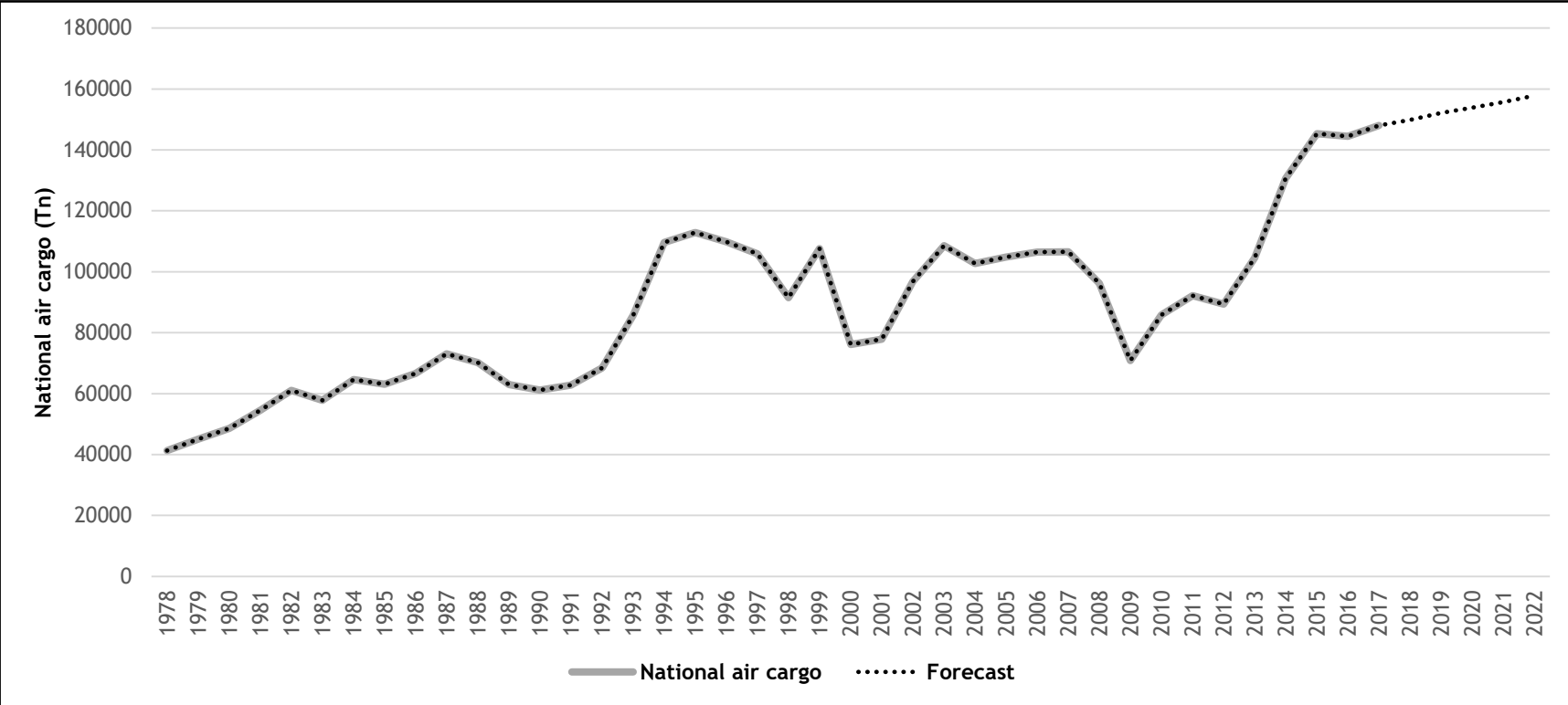
Table 3. Comparison of MAPE values for Model 3. Source: Authors.

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 a per capita and population were used as auxiliary variables to estimate the future forecast, thereby obtaining the results shown in Model 4.



- ▶ In Fig. 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 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.

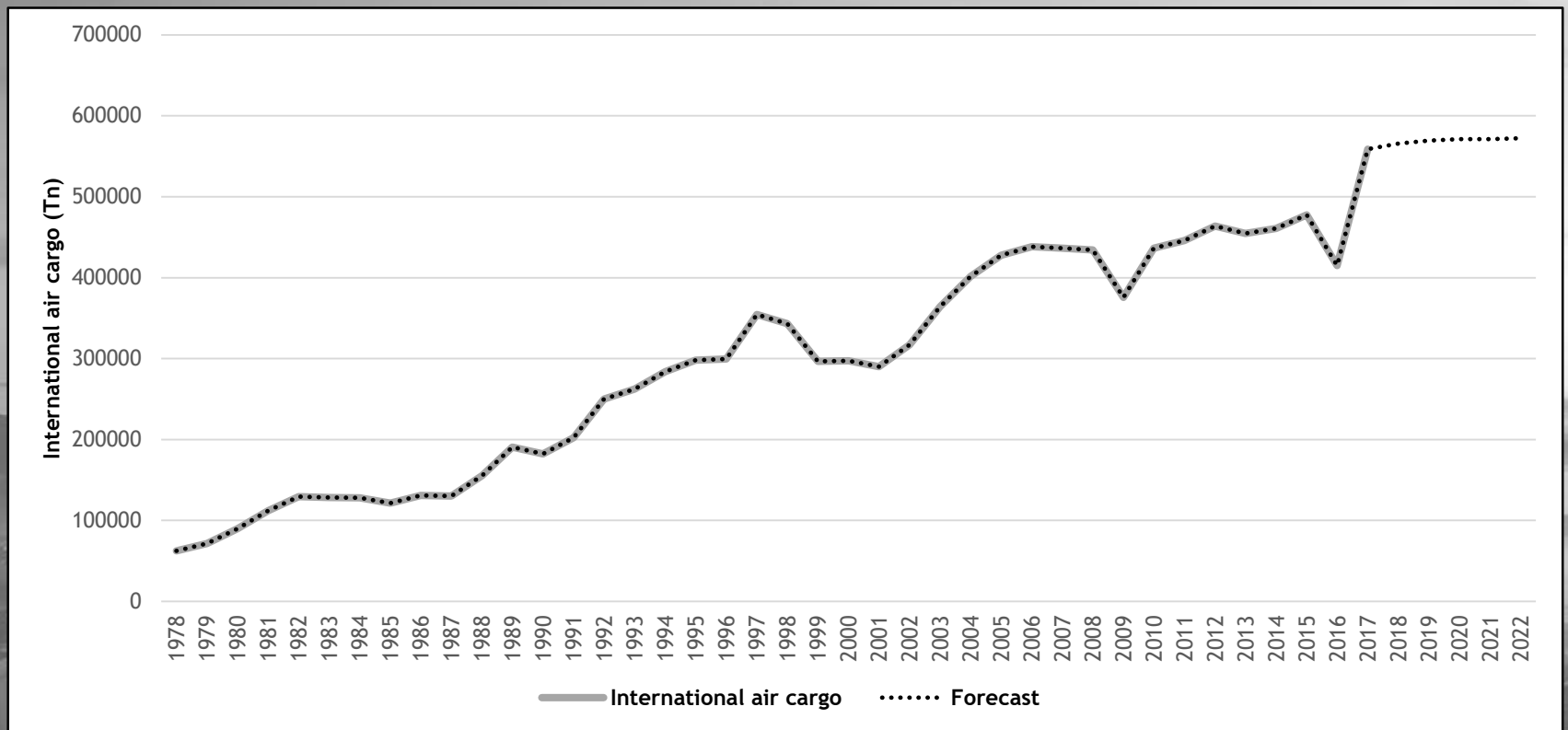
Table 4. Comparison of MAPE values for Model 4. Source: Authors.

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





► In the case of the variable “international 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 Fig.5).



- ▶ In Fig. 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) 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 5. Comparison of MAPE values for Model 5. Source: Authors.

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, imports, exports
4	0,00840434	International air cargo, GDP, imports
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





Conclusions

- ▶ Considering the advantages of using DLMs in the forecast of time series, an initial description of variables was made, which revealed a growing behavior as well as strong correlations in time with the covariables. As regards the covariables present in the models, an ARIMA model was used to carry out their future forecast and for those values to be included in the model chosen. The result of the application of DLMs presents MAPE values below 1%, which ensures high predictability forecasts.





- ▶ On the other hand, 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 the best performance as alternative models to develop reliable forecasts in air transport (or air traffic prognosis), at least in the short term.





- [1] Díaz Olariaga, O.; Ávila Álvarez, J. (2015). Evolution of the airport and air transport industry in Colombia and its impact on the economy. *Journal of Airline and Airport Management*, 5(1), 39-66.
- [2] Díaz Olariaga, O. (2017). Políticas de privatización de aeropuertos. El caso de Colombia. *Documentos y Aportes en Administración Pública y Gestión Estatal*, 29, 7-35.
- [3] Díaz Olariaga, O.; Zea, J.F. (2018). Influence of the liberalization of the air transport industry on configuration of the traffic in the airport network. *Transportation Research Procedia*, 33, 43-50.
- [4] Díaz Olariaga, O. (2018). Prognosis de tráfico aéreo en contexto de post-liberalización del transporte aéreo. El caso de Colombia. *Working Paper*. DOI: 10.13140/RG.2.2.11443.73760
- [5] West, M.; Harrison, J. (2006). *Bayesian forecasting and dynamic models*. New York: Springer Science & Business Media.
- [6] Honjo, K.; Shiraki, H.; Ashina, S. (2018). Dynamic linear modeling of monthly electricity demand in Japan: Time variation of electricity conservation effect. *PloS one*, 13(4), e0196331.





- [7] Petris, G.; Petrone, S., Campagnoli, P. (2009). *Dynamic Linear Models with R*. New York: Springer.
- [8] O'Connor, K. (2003). Global air travel: toward concentration or dispersal? *Journal of Transport Geography*, 11(2), 83-92.
- [9] Goetz, A. (2002). Deregulation, competition, and antitrust implications in the US airline industry. *Journal of Transport Geography*, 10(1), 1-19.
- [10] O'Kelly, M. (1998). A geographer's analysis of hub-and-spoke networks. *Journal of Transport Geography*, 6(3), 171-186.
- [11] Vowles, T. (2006). Airfare pricing determinants in hub-to-hub markets. *Journal of Transport Geography*, 14(1), 15-22.
- [12] Fan, T.; Vigeant-Langlois, L.; Geissler, C.; Bosler, B.; Wilmaking, J. (2001). Evolution of global airline strategic alliance and consolidation in the twenty-first century. *Journal of Air Transport Management*, 7(6), 349-360.
- [13] Tascón Hoyos, D.; Díaz Olariaga, O. (2018). Analysis of behavior of air passenger demand through system dynamics. 21st EURO Working Group on Transportation Meeting, 17-19 September 2018, Braunschweig, Germany.





Thank you!