Periodica Polytechnica Transportation Engineering, 52(1), pp. 75–85, 2024

Air Traffic Demand Forecasting with a Bayesian Structural Time Series Approach

Yesid Rodríguez[1], Oscar Díaz Olariaga[2*]

[1] Escuela de Posgrado, Universidad Konrad Lorenz, Carrera 9 Bis # 62-43, Bogotá, Colombia
[2] Facultad de Ingeniería Civil, Universidad Santo Tomás, Carrera 9 # 51-11, Bogotá, Colombia
[*] Corresponding author, e-mail: oscardiazolariaga@usta.edu.co

Received: 08 August 2022, Accepted: 10 October 2023, Published online: 17 October 2023

Abstract

Airport planning, and therefore the development of air infrastructure, depends to a large extent on the demand forecast for the future. To plan investments in the infrastructure of an airport system and to be able to meet future needs, it is essential to predict the level and distribution of demand, both for passengers and air cargo. In the present work, a forecast was made, in the medium-long term (10 years), of the demand for passengers and air cargo, applied to a specific case study (Colombia), and where the impact on air traffic during the most severe period of the COVID-19 pandemic was taken into account. To achieve this objective, and as a methodological approach, a model of the Bayesian Structural Time Series (BSTS) type was developed, designed to work with time series data, and widely used for feature selection, time series forecasting, and the immediate inference of the causal impact. From the results obtained, two relevant aspects can be highlighted, firstly, both demand and its growth trend will recover very soon (in just a couple of years), compared to the pre-pandemic year 2019, which was analyzed in the case study. And, secondly, the model presents very acceptable MAPE values (between 1% and 7%, depending on the variable to be forecasted), which makes the BSTS method a viable alternative methodology for calculating air traffic forecasts.

Keywords

airport, air transport, air traffic, aviation, forecast, Bayesian Structural Time Series

1 Introduction

The plans for the development of the different components of the airport system depend to a large extent on the levels of activity that are foreseen for the future. To plan the facilities and infrastructure of an airport, or system/group of airports, and to be able to meet future needs, it is essential to predict the level and distribution of demand in the various components of the airport system (TRB, 2002). Forecasting demand in an industry as dynamic and sensitive to exogenous factors as aviation is an extremely difficult task. However, it is necessary to carry out air traffic estimates as a preliminary step to the planning and design of airport facilities, whether of an airport or an airport system or network (Horonjeff et al., 2010; Wells and Young, 2004).

Understanding future demand patterns allows the airport planner to assess the future performance of the airport and thereby recommend consistent development programs to estimate the costs associated with these development plans and to project the sources and level of revenue to support the investments of capital to be realized (Ashford et al., 2011). Demand forecasting is a basic requirement to develop an airport master plan or an airport system plan at the regional or national level, thereby understanding the entire airport network of a region or country (García Cruzado, 2013; Horonjeff et al., 2010; ICAO, 1987, 2006; Janic, 2009; Janić, 2021; Rodriguez et al., 2020).

Finally, practical experience shows that air traffic forecasts are not usually accurate when the forecasts are made in the long term (20 years) or very long term (30 years), but the forecasts in the very short, short, medium, and medium-long-term are inevitably important for the planner and/or decision maker since they comprise a usual period of airport planning (or of an airport system/network) (ACI, 2016; de Neufville et al., 2013; Kazda and Caves, 2015).

Therefore, the objective of this research is to make a medium-long term forecast (10 years) of the demand for air passengers and air cargo (domestic and international for both variables), for which Colombia has been used as a case study (entire network of airports as a whole), with in a period of available monthly traffic data: 1992–2021, with the special feature of including demand data for

Cite this article as: Rodríguez, Y., Díaz Olariaga, O. (2024) "Air Traffic Demand Forecasting with a Bayesian Structural Time Series Approach", Periodica Polytechnica Transportation Engineering, 52(1), pp. 75–85. https://doi.org/10.3311/PPtr.20973

the years 2020 (which has been severely affected by the COVID-19 pandemic) and 2021 (pandemic transition and recovery period) in the analysis. To achieve this objective, and as a methodological technique, a model of the Bayesian Structural Time Series (BSTS) type was developed, which has significant advantages over classic models such as the SARIMA model, for example, Jalali and Rabotyagov (2020) mention certain advantages of state-space models such as the BSTS, in the first place they show that all ARIMA models can be expressed as a form of state-space, also that overtraining problem is solved concerning the other models when there is a limited number of observations over time, and can assign uncertainty parameters which give you an advantage when forecasting over time.

2 Literature review

Regarding the development of forecast for air transport/traffic, academics have been presenting formal research, with different methodologies and approaches, for approximately three decades. Now, it should be noted that despite its great importance, the analysis or research of air cargo transport demand forecasting is not present in so many publications, or at least not as much as the air passenger demand forecast (Baier et al., 2022). Thus, a variety of models have been developed to forecast air traffic demand (passengers and/or air cargo), many of the most widely used forecasting methods can be grouped into two large groups: economic models and time series models (Dantas et al., 2017). Economic methods focus on the correlation between passenger demand and multiple variables, such as the influence of changes in the economic environment and the traffic system, and then forecast models are established by a series of equations. Commonly used models include regression analysis (Abed et al., 2001), causality test (Fernandes and Pacheco, 2010), logit model (Garrow and Koppelman, 2004), and gravity model (Grosche et al., 2007). Time series methods rely primarily on historical data to predict by extracting the intrinsic relationship between the current data and the series of past observations. Various time series models have been used to forecast air passenger demand, such as smoothing techniques (Samagaio and Wolters, 2010), the adapted Markov model (Chin and Tay, 2001), ARIMA/SARIMA (Tsui et al., 2014), seasonal adjustment method (Aston and Koopman, 2006), etc.

However, due to the non-linear characteristic of air transport demand, economic and time series approaches are severely criticized due to their very limited and inefficient forecasting capacity (Tsui et al., 2014). For this reason, some academics try to explore other methodologies for forecasting air traffic demand, such as Dynamic Linear Models (DLM) (Rodriguez et al., 2020), which, compared to the usual forecast calculation methodologies, have the following advantages: detection of trends' stochastics that are hidden in the time series (West and Harrison, 2006) as well as the detection of structural changes that allow estimating the time-varying effect of exogenous shocks without increasing the number of parameters (Honjo et al., 2018); additionally, the conditional independence structure on which the dynamics of the state is based allows us to consider predictions considering a recursive algorithm (Petris et al., 2009). Another approach that has been used to forecast air passenger demand is System Dynamics (Suryani et al., 2012, 2010; Tascón and Díaz Olariaga, 2021); System Dynamics incorporates a set of tools to understand complex environments, and uses tools such as causality diagrams, fostering systemic thinking among process managers, captures the dynamic complexity of a given system, and provides the considerable advantage of simulating the model to evaluate variable interaction outcomes and anticipate long-term side effects of policies before they are implemented.

On the other hand, methodologies based on artificial neural networks (ANN) have been used for air traffic forecasting for several years. Several distinctive features of ANNs make them feasible and/or convenient for air traffic demand forecasting calculation, namely (Bao et al., 2012; Dingari et al., 2019; Gupta et al., 2019; Mostafaeipour et al., 2018; Srisaeng et al., 2015; Zhang et al., 1998):

- unlike traditional methods, ANNs are self-adaptive methods based on data in the sense that there are few a priori assumptions about the models for the problems under study. They learn from examples and capture subtle functional relationships between data, even if the underlying relationships are unknown or difficult to describe;
- ANNs are universal functional approximators, it has been shown that a network can approximate any continuous function with any desired precision. ANNs have more general and flexible functional forms (than traditional statistical methods) that they can handle effectively;
- finally, ANNs are not linear, that is, they are approaches based on nonlinear data, and capable of performing nonlinear models without prior knowledge about the relationships between input and output variables.

And finally, regarding the methodology used in this article for the development of air traffic demand, BSTS, very little is found in the scientific literature. Bazzo Vieira et al. (2022) used a BSTS model to estimate the impact of COVID-19 on daily trends in demand and emissions from commercial air travel (applied to the Brazilian air transport system). According to the authors, a BSTS model was chosen because, compared to other time series and difference-in-differences methods, it considers time-varying influences and confounding factors (allowing, for example, to isolate the unique impact of COVID-19). Al-Sultan et al. (2021) uses a wide range of time series forecast models, including the autoregressive integrated moving average (ARIMA) model, exponential smoothing with error term (ETS), Holt-Winters exponential smoothing, autoregression of neural network, and BSTS, to forecast in a very short-term air passenger demand at Kuwait Airport. The authors compare the performance of these models using the mean absolute percentage error (MAPE). According to the results, the BSTS model presents a better performance than other time series models in its ability to forecast complex time series; according to the authors, the BSTS model can be applied to study complex problems of irregular time series forecasting. Madhavan et al. (2020) forecast the air passenger and cargo demand of the Indian aviation industry using Autoregressive Integrated Moving Average (ARIMA) and BSTS models. The authors used 10-year (2009-2018) air cargo and passenger data. The study evaluated the ability of the ARIMA and BSTS models to incorporate uncertainty under dynamic adjustments. The findings inferred that, together with ARIMA, BSTS is well suited for the very short-term forecast of the four typical indicators of commercial air transport (international passengers, domestic passengers, international air cargo and domestic air cargo). Han and DeLaurentis (2011) developed a forecast model, based on Bayesian models, to forecast air traffic demand based on several scenarios that not only have predictive capacity but also consider constraints, and applied to the case of Chicago O'Hare Intl. Airport (USA). Based on the results, the authors state that the approach used, Bayesian models, shows, on the one hand, an acceptable ability to represent interactions between factors, and on the other hand, shows how predictive factors affect air traffic demand and considers restrictions, such as airport capacity. Finally, Xu et al. (2005) used a Bayesian model to forecast commercial flight delays (between airports). The results demonstrate higher predictive accuracy than a linear regression or Bayesian analysis with non-informative priors.

3 Conceptual framework

The modeling of variables in time has been a topic worked on by multiple academic branches, especially the space-state models (SSM) that were initially presented by control engineers (from Kalman (1960)), who found a wide application in processes that need continuous updating (Jones, 2019). These models are increasingly used in solving problems with variables of temporal type, including parameter estimation, smoothing, and predictions (Hamilton, 1994). The use of structural time series models can be classified as space models (state for time series data, which have advantages such as their flexibility since models such as ARIMA or VARMA can be expressed as a space model). In addition, they are modular models, since the model can be built from several state component models that capture the most important characteristics of the data, for example, the state components are used in capturing the trend, seasonality, among others (Koller and Friedman, 2009).

Based on the work of Scott and Varian (2014, 2015), the development of the BSTS model was deepened and extended, which can be used in the selection and short- and long-term forecast of time series. On the other hand, Brodersen et al. (2015) and Peters et al. (2017) present the inference of causal relationships of the model. Complementing the previous works, one of the pillars of the model is that it uses the Kalman Filter to be able to identify the non-measurable state of the linear dynamic system of the model (Durbin and Koopman, 2002; Petris et al., 2009). After the Kalman Filter, the selection of the 'spike and slab' variable (George and McCulloch, 1997) is taken into account, with which the regression predictors are selected throughout the modeling. Finally, the calculation of the average (Hoeting et al., 1999) is taken into account within the Bayesian model, which presents the combination of the results of the selection of characteristics and the prediction calculation. The use of the above allows to have a natural Bayesian interpretation and its presence in a BSTS model will allow for obtaining better results (Bach et al., 2013; Griffiths, 2003; Harvey et al., 2007).

According to Santana Jiménez (2020), BSTS models are structured under two components, on the one hand, a time series module that captures trend or seasonality patterns in the data. On the other hand, a regression element that ends up including potential exogenous data to reduce forecast error. Regarding the methodology for the development of this model, there is the Kalman filter, an algorithm that integrates the information provided to estimate the value of the target variables.

According to Zhang and Fricker (2021), BSTS is useful for time series forecasting, since it estimates the uncertainty in the predictions and the inference of their causal impact, from the use of credible Bayesian intervals, which synthesize the uncertainty of the inference. Among the advantages of this type of modeling are, on the one hand, that it provides parameters for the correct inclusion of the variables, and on the other hand, it offers the alternative of including the priority and weight of the model variables. Furthermore, Cerri et al. (2022) point out that this model can estimate the causal effect for a single target variable, generating a forecast for its future values based on a synthetic control, composed of untreated time series that were predictive of the target study variable at the time.

Finally, Giri et al. (2020) highlight the flexibility and adaptability that the BSTS has to adjust to the needs of an objective analysis, since this model allows decision-making, in the timing of the predictions (short or long term), on how to include the variables regressors and whether the data contain seasonality.

4 Application case

The data for the development of this research is obtained from the country-case of application (or study) Colombia, currently the third air market in the Latin American subcontinent, and fifth in the Americas, by volume of managed traffic (ACI, 2021; IATA, 2021). In Colombia, the air transport/aviation industry was liberalized In the early 1990s. This brought about structural reforms in both the airport and airline sectors, all through an uninterrupted battery of public policies (still in force today) that includes not only normative and regulatory aspects but also aggressive public and private investment programs in infrastructure and technology (Díaz Olariaga, 2021a).

Regarding the management of airport infrastructure, Colombia has followed the regional trend of concessioning the operation of said infrastructures to the private sector (Díaz Olariaga and Ávila Álvarez, 2015). Then, since the mid-1990s, and in various temporary phases, called generations, the Colombian government has granted various airports in the country, a total of 19, including the largest and most important (Díaz Olariaga and Pulido Moreno, 2019). As a result of public policies, both privatization and public and private investment was made in airport infrastructure, accompanied by policies of deregulation of the commercial aviation sector, where airfares have been fully liberalized since 2012 (Díaz Olariaga and Zea, 2018), From the beginning of the liberalization of the industry (1991) and until 2019, passenger transport (total) grew by almost 800% (Aerocivil, 2022; Díaz Olariaga, 2021b). On the other hand, in Colombia, the entry into the market of private air operators with a full-service business model, or FSC (Full-Service Carrier), occurred very soon after the liberalization of the sector (early-mid 1990s). But the entry into the market of low-cost airlines, or LCC (Low-Cost Carrier), was many years after liberalization (in 2012) (Díaz Olariaga and Zea, 2018).

In Colombia, as of March 20, 2020, all passenger air transport activities (domestic and international) were suspended, except in cases of humanitarian emergency. The transport of air cargo did not suffer this restriction and was able to continue its activity in an almost normal way (Díaz Olariaga and Alonso-Malaver, 2022). Air passenger traffic was reactivated in September 2020. This situation led to a precipitous drop in the demand for air passengers (85% in domestic passengers, 75% in international passengers, both compared to 2019); however, the demand for air cargo (total) only fell by 16% compared to 2019. By the beginning of 2021, almost all restrictions regarding the type of traffic (domestic / international) as its origin-destination had already been lifted. By the beginning of 2022, passenger demand was already showing a strong recovery trend (75% of domestic pax and 50% of international pax, both with respect to demand figures for 2019) (Aerocivil, 2022).

5 Methodology

5.1 Model

The Bayesian structural model is presented through the behavior of the time series using two equations, the observation equation and the transition equation (Scott and Varian, 2014). The observation equation is defined by Eq. (1):

$$y_t = \mu_t + \varepsilon_t, \tag{1}$$

with $t = 1, 2, \ldots, T$ and $\varepsilon_t \sim N(0, \sigma^2)$, what this equation denotes is that y_t is the value at the moment t of the time series which is related to a latent state called μ_t and finally the irregular component which is also called 'noise' or ε_t.

On the other hand, the transition equation is defined by Eq. (2):

$$\mu_{t+1} = \mu_t + \eta_t, \tag{2}$$

where $\eta_t \sim N(0, \sigma^2)$.

A model that can be written using these two equations (observation and transition) can be transformed into a basic structural model with three state components, a trend μ_t, a seasonal pattern τ_t and a regression component $\beta^T x_t$ (in this component are the socioeconomic covariates that are to be included in the model) (Eqs. (3), (4), (5) and (6)):

$$y_t = \mu_t + \tau_t + \beta^T x_t + \varepsilon_t , \tag{3}$$

where:

$$\mu_t = \mu_{t-1} + \delta_{t-1} + \eta_t , \tag{4}$$

$$\delta_t = \delta_{t-1} + w_t , \tag{5}$$

$$\tau_t = -\sum_{S-1}^{s-1} \tau_{t-s} + \gamma_t . \tag{6}$$

The variables ε, η, w, γ are normal random variables with constant variance over all time T, but different from each other; on the other hand:

- τ: random component S which is nothing more than the number of stations;
- δ_t: represents a random walk that determines the trend of the data;
- μ_t: represents the current trend, μ_t is affected by $\mu_{t-1} + \delta_{t-1}$ plus the noise generated by η_t;
- x_t: is the vector of socioeconomic covariates, represents the other time series that help to have a better estimate for y.

Among the reasons for using BSTS, it is found that it allows modeling a wide variety of components in time series such as trends, seasonality, cycles, etc. It allows to make models that adapt to the specific structure of the data. Furthermore, it uses Bayesian methods to estimate the model parameters. This means that you can incorporate previous information and update estimates as you collect more data, plus it provides Bayesian confidence intervals that reflect uncertainty in the forecasts. On the other hand, BSTS models are used to detect anomalies in time series data, this can be useful to identify outliers or unusual events (such as the COVID-19 pandemic). Finally, BSTS models allow the incorporation of covariates or exogenous variables into the model to improve the accuracy of predictions, which is a valuable capability when additional information is available that can help explain the time series, whereas ARIMA models they do not have a direct way of incorporating external covariates.

5.2 Data

For the present research, a time series of monthly traffic data is available from January 1992 to December 2021 (national and international passengers, and national and international air cargo of the Colombian air transport system as a whole, 58 airports open to commercial air traffic), data generated by the public aeronautical authority of Colombia (Aerocivil, 2022). On the other hand, there is also a set of socioeconomic variables (at the national level of the country-case study), namely: GDP, GDP per capita, Population, IPI (Industrial Production Index), CPI (Consumer Price Index), TRM (Representative Market Rate, or USD-COP exchange rate), FDI (Foreign Direct Investment), Exports, Imports, Gini Index, and Unemployment Rate, all of them for the entire study period (1992–2021), obtained from the public statistical agency of Colombia (DANE, 2022).

The pandemic generated an unexpected change in the habitual behavior of the series mentioned, for which an intervention analysis was carried out for April to October (of 2020) in the series of national passengers and for March to October (2020) in the international passenger series. Said intervention analysis consists of a compensated impulse that captures transient changes in level (Box and Tiao, 1975).

6 Results

For the proposed BSTS models, it is found that they present a mean absolute percentage error (MAPE), which measures the size of the (absolute) error in percentage terms, below 10%. The fact that the magnitude of the percentage error is estimated makes it an indicator frequently used by forecasters due to its easy interpretation. A small value of MAPE indicates that the forecasts have a higher probability of being accurate (Ren and Glasure, 2009).

In general terms, the use of socioeconomic covariates is presented according to the relationship that these can present from the economic point of view and how they could influence/impact the calculation of the forecast of each of the air transport variables to be predicted (Adenigbo et al., 2022; Kiracı and Battal, 2018; Navarro Hudiel and Acuña Mendoza, 2021; Rodriguez-Brindis et al., 2015), however, the effects of the COVID-19 pandemic (especially in the year 2020) on all socioeconomic variables (as they are included as covariates in the models) do not show a positive effect on the recovery trends that air transport has had, which is evidenced by presenting lower MAPE with respect to models that do not include such covariates. Therefore, the use of the BSTS model without covariates is the best prediction alternative for the period of time analyzed. In addition, given the structural behavior of the national and international passenger variables and national and international air cargo, it was shown that they present a repetitive or seasonal structure every 12 months, for which reason this component must be included in the BSTS model when the forecast estimate is made.

For the case of national (or domestic) passenger demand in the BSTS model, the MCMC algorithm (Hoeting et al., 1999; Madigan and Raftery, 1994) with 1000 iterations was used to sample from the posterior distribution, with a seasonal component of order 12, a MAPE of 4.69% was obtained, resulting in that the estimated or projected demand of domestic passengers for the year 2031 will be 54.4 million (Fig. 1).

Table 1 presents the results of the MAPE of the BSTS models for the forecast of the domestic passenger variable without covariates and including three covariates (socioeconomic) individually. The results show that the second best model was achieved with the CPI (Consumer Price Index) variable, a model in which its best MAPE (8.10%) was found using 1,000 iterations and a seasonal component of order 12, following the order the following model was found with the variable TRM (Representative Market Rate or USD-COP exchange rate), a model in which its best MAPE (11.19%) was found using 10,000 iterations and a seasonal component of order 12 and finally, the model with the Unemployment Rate variable, a model in which its best MAPE (12.95%) was found using 10,000 iterations and a seasonal component of order 12.

In the case of international passenger demand in the BSTS model, to make a sample from the posterior distribution, the MCMC algorithm is used with 1,000 iterations, with a seasonal component of order 12, a MAPE of 7.4% was obtained, resulting in that the expected demand for international passengers for the year 2031 will be 22.2 million (Fig. 2).

Table 2 presents the results of the MAPE of the BSTS models for the forecast of the international passenger variable without covariates and including three covariates (socioeconomic) individually. The results show that the second best model was achieved with the variable TRM (Representative Market Rate or USD-COP exchange rate), a model in which its best MAPE (8.75%) was found using 10,000 iterations and a seasonal component of order 12, following the order, the following model was found with the variable CPI (Consumer Price Index), a model in which its best MAPE (9.42%) was found using 5,000 iterations and a seasonal component of order 12, and finally, the model with the Unemployment Rate, a model in which its best MAPE (10.42%) was found using 1,000 iterations and a seasonal component of order 12.

For the case of the domestic air cargo demand in the BSTS model, to make a sample from the posterior distribution, the MCMC algorithm was used with 2,000 iterations, with a seasonal component of order 24, a MAPE of 1.3% was obtained, resulting in that the expected demand for domestic air cargo transport for the year 2031 will be 318,000 tons (Fig. 3).

Table 3 presents the results of the MAPE of the BSTS models for the forecast of domestic air cargo without covariates and including four covariates (socioeconomic)

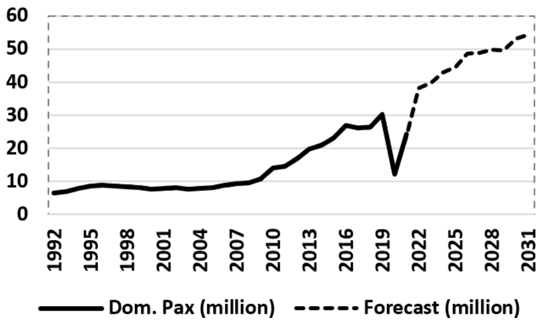

Fig. 1 Traffic history (1992–2021) and forecast (2022–2031) of domestic air passenger demand. Source: authors

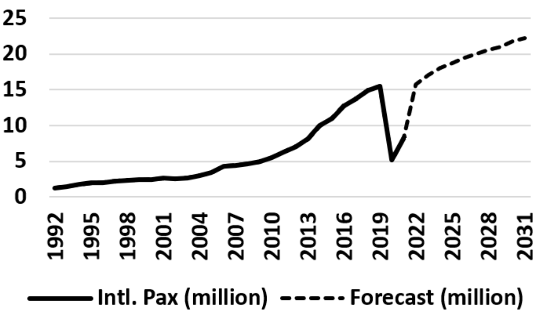

Fig. 2 Traffic history (1992–2021) and forecast (2022–2031) of international air passenger demand. Source: authors

Table 1 Comparison of MAPE for models with covariates (socioeconomic) for domestic passenger demand. Source: authors

Model	Variable and covariate	MAPE
Model 1	Domestic Pax (no covariates)	4.69%
Model 2	Domestic Pax (with CPI)	8.10%
Model 3	Domestic Pax (with TRM)	11.19%
Model 4	Domestic Pax (with Unemployment Rate)	12.95%

Table 2 Comparison of MAPE for models with covariates (socioeconomic) for international passenger demand. Source: authors

Model	Variable and covariate	MAPE
Model 1	Intl. Pax (no covariates)	7.41%
Model 2	Intl. Pax (with TRM)	8.75%
Model 3	Intl. Pax (with CPI)	9.42%
Model 4	Intl. Pax (with Unemployment Rate)	10.42%

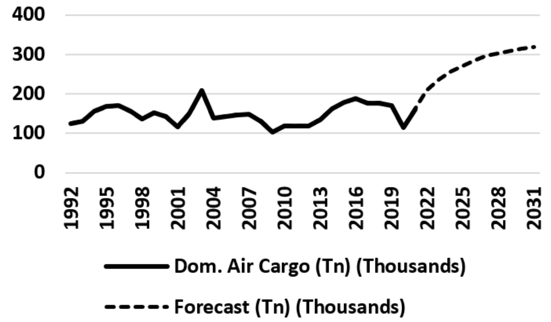

Fig. 3 Traffic history (1992–2021) and forecast (2022–2031) of domestic air cargo demand. Source: authors

Table 3 Comparison of MAPE for models with covariates (socio-economic) for domestic air cargo transport demand. Source: authors

Model	Variable and covariate	MAPE
Model 1	Dom. Air Cargo (no covariates)	1.31%
Model 2	Dom. Air Cargo (with TRM)	6.48%
Model 3	Dom. Air Cargo (with Imports)	8.07%
Model 4	Dom. Air Cargo (with Exports)	8.31%

individually. The results show that the second best model was achieved with the variable TRM (Representative Market Rate or USD-COP exchange rate), a model in which its best MAPE (6.48%) was found using 2,000 iterations and a seasonal component of order 12, the next best model was found with the variable Imports, a model in which its best MAPE (6.48%) was found using 10,000 iterations and a seasonal component of order 12, and the last model is with the variable Exports in which its best MAPE (8.31%) was found using 2,000 iterations and a seasonal component of order 12.

For the case of international air cargo demand in the BSTS model, to sample from the posterior distribution, the MCMC algorithm was used with 5,000 iterations, with a seasonal component of order 12, obtaining a MAPE of 2.21%, obtaining that the expected demand for international air cargo transport for the year 2031 will be 769,700 tons (Fig. 4).

Table 4 presents the results of the MAPE of the BSTS models for the forecast of the international air cargo demand variable without covariates and including four covariates (socioeconomic) individually. The results show that the second-best model was achieved with the variable TRM (Representative Market Rate or USD-COP exchange rate), a model in which its best MAPE (4.08%) was found using 5,000 iterations and a seasonal component. of order 12; the following model was found with the variable Exports, a model in which its best MAPE (7.98%) was found using 20,000 iterations and a seasonal component of order 12,

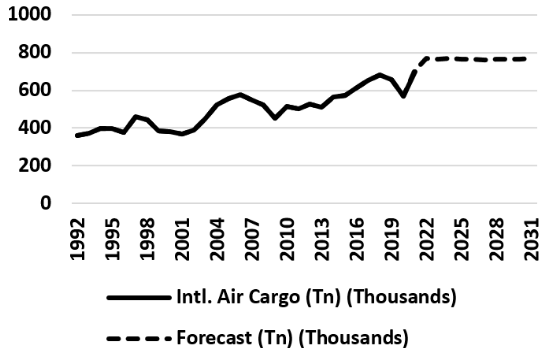

Fig. 4 Traffic history (1992–2021) and forecast (2022–2031) of international air cargo demand. Source: authors

Table 4 Comparison of MAPE for models with covariates (socio-economic) for international air cargo transport demand. Source: authors

Model	Variable and covariate	MAPE
Model 1	Intl. Air Cargo (no covariates)	2.21%
Model 2	Intl. Air Cargo (with TRM)	4.08%
Model 3	Intl. Air Cargo (with Exports)	7.98%
Model 4	Intl. Air Cargo (with Imports)	11.95%

and finally, the model with the variable Imports generated a MAPE of 11.95% and was achieved using 1,000 iterations and a seasonal component of order 12.

It should be mentioned that with the variables used in this investigation it is not possible to improve the values obtained from MAPE (already very acceptable), the MAPE could be improved by increasing the number of variables, but from the country-case study used here no more than those already mentioned above are available (which are also those commonly used in all forecast studies).

In another order, highlighting that the advantages of using MAPE consist, first of all, in its interpretation, since it is expressed as a percentage, which makes its interpretation intuitive and easy to understand, and it is also independent of the scale of data, which means that it can be used to compare the accuracy of forecasts in different contexts, making it easy to compare between different models or data sets. And secondly, the MAPE proportionally penalizes the errors, which means that large errors have a greater impact on the metric than small errors, this adequately reflects the importance of errors in the forecast and is especially useful when you want to give priority to accuracy in large forecasts.

Finally, it is worth mentioning that the results obtained here coincide with the forecasts made in a recent study (Gudmundsson et al., 2021), on the recovery of air transport worldwide, which predicts that the demand for air passengers will recover, to pre-pandemic levels, by the end of 2022 (optimistic scenario). And the results also coincide with very recent studies and projections of the three major international air transport organizations. For example, IATA (2022) estimates a recovery of both the absolute value of passenger demand and the growth trend (at the pre-pandemic rate) by 2023 for the Latin America & Caribbean (LA&C), region where the country-case is located (of the present research); ICAO (2022) presents similar passenger demand projections for LA&C, although the demand growth trend at 2019 values is at the end of 2022; and finally ACI (2022) predicts that passenger demand, worldwide, will recover the values of 2019 at the end of 2023, for domestic passengers, and in mid-2024 for international passengers.

7 Conclusions

Firstly, estimating the demand for air transport (passenger and cargo) at the local/national level provides valuable information so that civil aviation planners can design, plan and implement, well in advance: air infrastructure development strategies; a capital investment schedule (public and/or private); and finally, related public policies that allow decision makers (public) to contribute to consolidate and strengthen the development of the local/national air transport industry. Secondly, both academia and industry, and at a global level, are developing research and forecasting studies that allow them to know when and how air traffic (both demand volume and growth trend) will recover to pre-pandemic levels (the year 2019), due to the importance of the aviation industry, not only in terms of local/regional/global connectivity but also its contribution to the global economy.

Regarding the methodology used in this research, and the results obtained, it can be highlighted that the forecasts are in line with the behavior of the post-pandemic passenger and air cargo series, showing the future growth of the aviation sector. In addition, it is evident that the use of the BSTS models without covariates and with covariates present acceptable results in the forecast, which were justified with the MAPE showing values below 10% error, they are also models that present robustness when it comes to model series that have had shocks in their structure, such as the one that occurred due to the global health emergency of COVID-19 and that brought the values of the series to minimum levels and then began to return to their usual state. The use of Bayesian models opens the possibility of working with long-term aeronautical time series forecasts, presenting an advantage over classic time series models that converge to the mean in the long term, in addition to not assimilating changes in the structure of the series as if the models presented do.

For continuity to future related studies, several alternatives are presented, namely:

1. although the historical data used is extensive (1992–2021 and monthly), this range could be extended, for example, ten or fifteen more years, to see how the model (BSTS) behaves with a greater amount of data, although this represents a challenge (since in air transport it is not always possible to obtain such old and quality data);
2. although the current model presents an excellent performance for a medium-term forecast (10 years), it would be interesting to see how the model behaves for a longer forecast horizon, for example at 20 and 30 years (usual forecast times for the demand in the master plans of the airports, for their planning in the long and very long term);
3. considering that the COVID-19 pandemic had a strong impact on the drop in demand, it would be interesting to use, simultaneously with the BSTS model, other methodologies to compare the behavior and performance of the different models in relation to the impact produced by the COVID-19 pandemic in traffic data;
4. and, finally, and using BSTS models, expand the traffic demand forecast at a geographical level, for example, including some countries (the largest) in the Latin American region (where Colombia is located, country-case study), to find out and compare the recovery trends in traffic demand for the post-pandemic period.

References

Abed, S. Y., Ba-Fail, A. O., Jasimuddin, S. M. (2001) "An econometric analysis of international air travel demand in Saudi Arabia", Journal of Air Transport Management, 7(3), pp. 143–148. https://doi.org/10.1016/S0969-6997(00)00043-0

ACI (2016) "Traffic Forecast", Airports Council International, Montreal, Canada, Technical paper.

ACI (2021) "Annual World Airport Traffic Report", Airports Council International, Montreal, Canada.

ACI (2022) "The impact of COVID-19 on airports - and the path to recovery", Airports Council International World, Oct. 6. [online] Available at: https://aci.aero/2022/10/06/the-impact-of-covid-19-on-airports-and-the-path-to-recovery/ [Accessed: 09 October 2023]

Adenigbo, A. J., Mageto, J., Luke, R. (2022) "Macroeconomic Determinants of Air Cargo Flows in Ghana", Latin American Journal of Trade Policy, 5(12), pp. 7–36. https://doi.org/10.5354/0719-9368.2022.67061

Aerocivil (2022) "Estadísticas (Statistics)", [online] Available at: https://www.aerocivil.gov.co/ [Accessed: 07 August 2022] (in Spanish)

Al-Sultan, A. T., Al-Rubkhi, A., Alsaber, A., Pan, J. (2021) "Forecasting air passenger traffic volume: evaluating time series models in long-term forecasting of Kuwait air passenger data", Advances and Applications in Statistics, 70(1), pp. 69–89. https://doi.org/10.17654/AS070010069

Ashford, N. J., Mumayiz, S., Wright, P. H. (2011) "Airport Engineering: Planning, Design, and Development of 21st Century Airports", John Wiley & Sons. ISBN 9780470398555 https://doi.org/10.1002/9780470950074

Aston, J. A. D., Koopman, S. J. (2006) "A Non-Gaussian Generalization of the Airline Model for Robust Seasonal Adjustment", Journal of Forecasting, 25(5), pp. 325–349. https://doi.org/10.1002/for.991

Bach, S., Huang, B., London, B., Getoor, L. (2013) "Hinge-loss Markov Random Fields: Convex Inference for Structured Prediction", [preprint] arXiv:1309.6813, 26 September 2013. https://doi.org/10.48550/arXiv.1309.6813

Baier, F., Berster, P., Gelhausen, M. (2022) "Global cargo gravitation model: airports matter for forecasts", International Economics and Economic Policy, 19(1), pp. 219–238. https://doi.org/10.1007/s10368-021-00525-2

Bao, Y., Xiong, T., Hu, Z. (2012) "Forecasting Air Passenger Traffic by Support Vector Machines with Ensemble Empirical Mode Decomposition and Slope-Based Method", Discrete Dynamics in Nature and Society, 2012, 431512. https://doi.org/10.1155/2012/431512

Bazzo Vieira, J. P., Vieira Braga, C. K., Pereira, R. H. M. (2022) "The impact of COVID-19 on air passenger demand and CO_2 emissions in Brazil", Energy Policy, 164, 112906. https://doi.org/10.1016/j.enpol.2022.112906

Box, G. E. P., Tiao, G. C. (1975) "Intervention analysis with applications to economic and environmental problems", Journal of the American Statistical Association, 70(349), pp. 70–79. https://doi.org/10.1080/01621459.1975.10480264

Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., Scott, S. L. (2015) "Inferring causal impact using Bayesian structural time-series models", The Annals of Applied Statistics, 9(1), pp. 247–274. https://doi.org/10.1214/14-AOAS788

Cerri, J., Carnevali, L., Monaco, A., Genovesi, P., Bertolino, S. (2022) "Blacklists do not necessarily make people curious about invasive alien species. A case study with Bayesian structural time series and Wikipedia searches about invasive mammals in Italy", NeoBiota, 71, pp. 113–128. https://doi.org/10.3897/neobiota.71.69422

Chin, A. T. H., Tay, J. H. (2001) "Developments in air transport: implications on investment decisions, profitability and survival of Asian airlines", Journal of Air Transport Management, 7(5), pp. 319–330. https://doi.org/10.1016/S0969-6997(01)00026-6

DANE (2022) "Departamento Administrativo Nacional de Estadística (Colombia)" (National Administrative Department of Statistics (Colombia)), [online] Available at: https://www.dane.gov.co/ [Accessed: 07 August 2022] (in Spanish)

Dantas, T. M., Cyrino Oliveira, F. L., Varela Repolho, H. M. (2017) "Air transportation demand forecast through Bagging Holt Winters methods", Journal of Air Transport Management, 59, pp. 116–123. https://doi.org/10.1016/j.jairtraman.2016.12.006

de Neufville, R., Odoni, A. R., Belobaba, P. P., Reynolds, T. G. (2013) "Airport Systems, Planning, Design, and Management", McGraw-Hill Education LLC. ISBN 9780071770583

Díaz Olariaga, O. (2021a) "Contribución del transporte aéreo a la conectividad territorial. El caso de Colombia" (Contribution of air transport to territorial connectivity. The case of Colombia), EURE, 47(140), pp. 117–141. (in Spanish) https://doi.org/10.7764/EURE.47.140.06

Díaz Olariaga, O. (2021b) "The Role of Regional Airports in Connectivity and Regional Development", Periodica Polytechnica Transportation Engineering, 49(4), pp. 394-406. https://doi.org/10.3311/PPtr.16557

Díaz Olariaga, O., Alonso-Malaver, C. (2022) "Impact of airport policies on regional development. Evidence from the Colombian case", Regional Science Policy & Practice, 14(6), pp. 185–210. https://doi.org/10.1111/rsp3.12483

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), pp. 39–66. https://doi.org/10.3926/jairm.43

Díaz Olariaga, O., Pulido Moreno, L. (2019) "Measurement of airport efficiency. The case of Colombia", Transport and Telecommunication, 20(1), pp. 40–51. https://doi.org/10.2478/ttj-2019-0004

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, pp. 43–50. https://doi.org/10.1016/j.trpro.2018.10.074

Dingari, M., Reddy, D. M., Sumalatha, V. (2019) "Air Traffic Forecasting Using Artificial Neural Networks", International Journal of Scientific & Technology Research, 8(10), pp. 556–559.

Durbin, J., Koopman, S. J. (2002) "A simple and efficient simulation smoother for state space time series analysis", Biometrika, 89(3), pp. 603–616. https://doi.org/10.1093/biomet/89.3.603

Fernandes, E., Pacheco, R. R. (2010) "The causal relationship between GDP and domestic air passenger traffic in Brazil", Transportation Planning and Technology, 33(7), pp. 569–581. https://doi.org/10.1080/03081060.2010.512217

García Cruzado, M. (2013) "Aeropuertos. Planificación, Diseño y Medio Ambiente" (Airports. Planning, Design and Environment), Ibergarceta Publicaciones. ISBN 9788415452799 (in Spanish)

Garrow, L. A., Koppelman, F. S. (2004) "Predicting air travelers' no-show and standby behavior using passenger and directional itinerary information", Journal of Air Transport Management, 10(6), pp. 401–411. https://doi.org/10.1016/j.jairtraman.2004.06.007

George, E. I., McCulloch, R. E. (1997) "Approaches for Bayesian Variable Selection", Statistica Sinica, 7(2), pp. 339–373.

Giri, S., Purkayastha, S., Hazra, S., Chanda, A., Das, I., Das, S. (2020) "Prediction of Monthly Hilsa (*Tenualosa ilisha*) Catch in the Northern Bay of Bengal using Bayesian Structural Time Series Model", Regional Studies in Marine Science, 39, 101456. https://doi.org/10.1016/j.rsma.2020.101456

Griffiths, W. E. (2003) "Bayesian Inference in the Seemingly Unrelated Regressions Model", CRC Press. ISBN 9780429213700

Grosche, T., Rothlauf, F., Heinzl, A. (2007) "Gravity models for airline passenger volume estimation", Journal of Air Transport Management, 13(4), pp. 175–183. https://doi.org/10.1016/j.jairtraman.2007.02.001

Gudmundsson, S. V., Cattaneo, M., Redondi, R. (2021) "Forecasting temporal world recovery in air transport markets in the presence of large economic shocks: The case of COVID-19", Journal of Air Transport Management, 91, 102007. https://doi.org/10.1016/j.jairtraman.2020.102007

Gupta, V., Sharma, K., Sangwan, M. S. (2019) "Airlines passenger forecasting using LSTM based recurrent neural networks", International Journal "Information Theories and Applications", 26(2), pp. 178–187.

Hamilton, J. D. (1994) "Time Series Analysis", Princeton University Press. ISBN 9780691042893

Han, S., DeLaurentis, D. (2011) "Air Traffic Demand Forecast at a Commercial Airport using Bayesian Networks", In: 11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Virginia Beach, VA, USA, pp. 1524–1536. ISBN 9781618393302

Harvey, A. C., Trimbur, T. M., Van Dijk, H. K. (2007) "Trends and Cycles in Economic Time Series: A Bayesian Approach", Journal of Econometrics, 140(2), pp. 618–649. https://doi.org/10.1016/j.jeconom.2006.07.006

Hoeting, J. A., Madigan, D., Raftery, A. E., Volinsky, C. T. (1999) "Bayesian Model Averaging: A Tutorial", Statistical Science, 14(4), pp. 382–401.

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. https://doi.org/10.1371/journal.pone.0196331

Horonjeff, R., McKelvey, F., Sproule, W. J., Young, S. B. (2010) "Planning and Design of Airports", McGraw-Hill. ISBN 978-0071446419

IATA (2021) "World Air Transport Statistics 2021", International Air Transport Association, Montreal, Canada.

IATA (2022) "Global Outlook for Air Transport: Sustained Recovery Amidst Strong Headwinds", International Air Transport Association, Geneva, Switzerland.

ICAO (1987) "Airport Planning Manual: Part 1: Master Planning", International Civil Aviation Organization, Montreal, Canada, Doc 9184-AN/902.

ICAO (2006) "Manual on Air Traffic Forecasting", International Civil Aviation Organization, Montreal, Canada, Doc 8991.

ICAO (2022) "Effects of Novel Coronavirus (COVID 19) on Civil Aviation: Economic Impact Analysis", [pdf] International Civil Aviation Organization, Montreal, Canada. Available at: https://www.slideshare.net/ssuser942e40/icao-covid19-economic-impact2022-06-10-1pdf [Accessed: 09 October 2023]

Jalali, P., Rabotyagov, S. (2020) "Quantifying cumulative effectiveness of green stormwater infrastructure in improving water quality", Science of the Total Environment, 731, 138953. https://doi.org/10.1016/j.scitotenv.2020.138953

Janic, M. (2009) "The airport analysis, planning and design: demand, capacity and congestion", Nova Science Publishers. ISBN 978-1-60741-3080

Janić, M. (2021) "System Analysis and Modelling in Air Transport: Demand, Capacity, Quality of Services, Economic, and Sustainability", CRC Press. ISBN 9780429321276 https://doi.org/10.1201/9780429321276

Jones, R. H. (2019) "Longitudinal Data with Serial Correlation: A state-space approach", Chapman & Hall. ISBN 9780367450083

Kalman, R. E. (1960) "A new approach to linear filtering and prediction problems", Journal of Basic Engineering, 82(1), pp. 35–45. https://doi.org/10.1115/1.3662552

Kazda, A., Caves, R. E. (2015) "Airport design and operation", Emerald Publishing. ISBN 9781784418700

Kiracı, K., Battal, Ü. (2018) "Macroeconomic Determinants of Air Transportation: A VAR Analysis on Turkey", Gaziantep University Journal of Social Sciences, 17(4), pp. 1536–1557. https://doi.org/10.21547/jss.391041

Koller, D., Friedman, N. (2009) "Probabilistic Graphical Models: Principles and Techniques", The MIT Press. ISBN 9780262013192

Madhavan, M., Sharafuddin, M. A., Piboonrungroj, P., Yang, C.-C. (2020) "Short-term forecasting for airline industry: the case of Indian air passenger and air cargo", Global Business Review. https://doi.org/10.1177/0972150920923316

Madigan, D., Raftery, A. E. (1994) "Model selection and accounting for model uncertainty in graphical models using Occam's window", Journal of the American Statistical Association, 89(428), pp. 1535–1546. https://doi.org/10.2307/2291017

Mostafaeipour, A., Goli, A., Qolipour, M. (2018) "Prediction of air travel demand using a hybrid artificial neural network (ANN) with Bat and Firefly algorithms: a case study", The Journal of Supercomputing, 74(10), pp. 5461–5484. https://doi.org/10.1007/s11227-018-2452-0

Navarro Hudiel, S. J., Acuña Mendoza, J. L. (2021) "Determinación de tasas de crecimiento de tráfico promedio diario anual en Nicaragua a partir de datos macroeconómicos" (Determination of the average daily annual traffic growth rates in Nicaragua based on macroeconomic data), Revista Ciencia y Tecnología El Higo, 11(2), pp. 70–83. (in Spanish) https://doi.org/10.5377/elhigo.v11i2.13033

Peters, J., Janzing, D., Schölkopf, B. (2017) "Elements of Causal Inference: Foundations and Learning Algorithms", The MIT Press. ISBN 9780262037310

Petris, G., Petrone, S., Campagnoli, P. (2009) "Dynamic linear models", In: Dynamic Linear Models with R, Springer, pp. 31–84. ISBN 978-0-387-77237-0 https://doi.org/10.1007/b135794_2

Ren, L., Glasure, Y. (2009) "Applicability of the revised mean absolute percentage errors (MAPE) approach to some popular normal and non-normal independent time series", International Advances in Economic Research, 15(4), pp. 409–420. https://doi.org/10.1007/s11294-009-9233-8

Rodriguez-Brindis, M. A., Mejia-Alzate, M. L., Zapata Aguirre, S. (2015) "La causalidad entre el crecimiento económico y la expansión del transporte aéreo: un análisis empírico para Chile" (The causality between economic growth and the expansion of air transport: an empirical analysis for Chile), Revista de Economía Del Rosario, 18(1), pp. 127–144. (in Spanish) https://doi.org/10.12804/rev.econ.rosario.18.01.2015.04

Rodriguez, Y., Pineda, W., Diaz Olariaga, O. (2020) "Air traffic forecast in post-liberalization context: a Dynamic Linear Models approach", Aviation, 24(1), pp. 10–19. https://doi.org/10.3846/aviation.2020.12273

Santana Jiménez, L. J. (2020) "Nowcasting con Google Trends: Dinámica de la Actividad Económica Mensual, el Consumo Privado y la Inversión basada en datos de Google Trends y un Modelo Bayesiano Estructural de Series de Tiempo" (Nowcasting with Google Trends: Dynamics of the Monthly Economic Activity, Private Consumption and Investment based on Google Trends Data and a Bayesian Structural Time Series Model), presented at XIII Foro de Investigadores de Bancos Centrales del Consejo Monetario Centroamericano, Ciudad de Guatemala, Guatemala, Sept. 5–6. [online] Available at: https://www.secmca.org/recard/index.php/foro/article/view/154 [Accessed: 09 October 2023] (in Spanish)

Samagaio, A., Wolters, M. (2010) "Comparative analysis of government forecasts for the Lisbon Airport", Journal of Air Transport Management, 16(4), pp. 213–217. https://doi.org/10.1016/j.jairtraman.2009.09.002

Scott, S. L., Varian, H. R. (2014) "Predicting the present with Bayesian structural time series", International Journal of Mathematical Modelling and Numerical Optimisation, 5(1–2), pp. 4–23. https://doi.org/10.1504/IJMMNO.2014.059942

Scott, S. L., Varian, H. R. (2015) "Bayesian Variable Selection for Nowcasting Economic Time Series", In: Goldfarb, A., Greenstein, S. M., Tucker, C. E. (eds.) Economic Analysis of the Digital Economy, NBER Chapters, National Bureau of Economic Research, Inc., pp. 119–135. ISBN 978-0-226-20684-4

Srisaeng, P., Baxter, G., Wild, G. (2015) "Using an artificial neural network approach to forecast Australia's domestic passenger air travel demand", World Review of Intermodal Transportation Research, 5(3), pp. 281–313. https://doi.org/10.1504/WRITR.2015.069243

Suryani, E., Chou, S.-Y., Chen, C.-H. (2010) "Air passenger demand forecasting and passenger terminal capacity expansion: A system dynamics framework", Expert Systems with Applications, 37(3), pp. 2324–2339. https://doi.org/10.1016/j.eswa.2009.07.041

Suryani, E., Chou, S.-Y., Chen, C.-H. (2012) "Dynamic simulation model of air cargo demand forecast and terminal capacity planning", Simulation Modelling Practice and Theory, 28, pp. 27–41. https://doi.org/10.1016/j.simpat.2012.05.012

Tascón, D. C., Díaz Olariaga, O. (2021) "Air traffic forecast and its impact on runway capacity. A System Dynamics approach", Journal of Air Transport Management, 90, 101946. https://doi.org/10.1016/j.jairtraman.2020.101946

TRB (2002) "Aviation Demand Forecast: Survey of Methodologies", Transportation Research Board, Washington, DC, USA, Rep. E-C040.

Tsui, W. H. K., Balli, H. O., Gilbey, A., Gow, H. (2014) "Forecasting of Hong Kong airport's passenger throughput", Tourism Manage, ment, 42, pp. 62–76. https://doi.org/10.1016/j.tourman.2013.10.008

Wells, A. T., Young, S. B. (2004) "Airport Planning & Management", McGraw-Hill. ISBN 9780071413015

West, M., Harrison, J. (2006) "Bayesian forecasting and dynamic models", Springer Science & Business Media. ISBN 9780387227771

Xu, N., Donohue, G., Laskey, K. B., Chen, C. H. (2005) "Estimation of Delay Propagation in Aviation System using Bayesian Network", presented at 6th USA-EUROPE ATM Seminar, Baltimore, MD, USA, Jun. 27–30.

Zhang, G., Patuwo, B. E., Hu, M. Y. (1998) "Forecasting with artificial neural networks: The state of the art", International Journal of Forecasting, 14(1), pp. 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7

Zhang, Y., Fricker, J. D. (2021) "Quantifying the Impact of COVID-19 on Non-Motorized Transportation: A Bayesian Structural Time Series Model", Transport Policy, 103, pp. 11–20. https://doi.org/10.1016/j.tranpol.2021.01.013