



DOI 10.2478/sbe-2019-0028
SBE no. 14(2) 2019

FORECASTING PASSENGER TRAFFIC FOR A REGIONAL AIRPORT

OPREANA ALIN

Lucian Blaga University of Sibiu, Romania

ȚICHINDELEAN MIHAI

Lucian Blaga University of Sibiu, Romania

MIHAIU DIANA MARIETA

Lucian Blaga University of Sibiu, Romania

TILEAGĂ COSMIN

Lucian Blaga University of Sibiu, Romania

Abstract:

The purpose of the present research is estimating the potential traffic for SIA (Sibiu International Airport, SBZ) for the year 2017. Predicting as accurate as possible the passenger traffic for a certain airport is an aspect of major importance for both the airport management and the airline companies. The theoretical quality of the forecasting models for air traffic of passengers is fundamental for obtaining the most accurate predictions. In this regard, a two-step process was used in developing the traffic forecasting model: (1) Identifying the proper regression model for traffic estimation based on the number of aircraft departures, and (2) Forecasting the number of aircraft departures for the current routes operated SIA. The predicted total passenger traffic overestimates the actual total traffic with only 2.4% and the actual total traffic without the transit traffic with only 1.42%.

Key words: *air traffic, traffic forecasting model, regression model, regional airport*

1. Introduction

Predicting as accurate as possible the passenger traffic for a certain airport is an aspect of major importance for both the airport management and the airline companies. The airport management must have quality forecasts regarding the future passenger flow in order to be able to properly decide regarding the future investments in the airport infrastructure, personnel policy and airport tariff policy. Moreover, airline companies use traffic forecasts to develop their strategy of operating new routes on certain destinations, to

adapt their flight frequencies for the destinations operated and to establish their price policy. Thus, it is essential that the model used for predicting passenger traffic generates accurate results because these are further used for optimal allocation of financial resources in various airport investments.

Bent Flyvbjerg, Mette K. Skamris Holm, Soren L. Buhl have observed a tendency of used forecast models that overestimate future passenger traffic (B. Flyvbjerg, M. K. S. Holm, S. L. Buhl, 2007). Opinions on the future evolution of air traffic are contrary. On the one hand, future air traffic could be negatively influenced by the development of high-speed rail infrastructure (P. Parthasarathi, D. Levinson, 2010), but on the other hand, the air services provided by low-cost carriers stimulate the air traffic. A recent study on the European market shows that the development of high-speed rail networks has led to a minor reduction in flight demand, but the expansion of low-cost carriers has generated a significant increase in air traffic (Regina R. Clewlow, Joseph M. Sussman, Hamsa Balakrishnan, 2014).

Contrary to these studies, the air transport passenger demand is still growing year by year. Statistical data published by ACI International shows a 7.5% worldwide increase in air passenger traffic in 2017 compared to the level of 2016. At the European level, ACI Europe reports a 8.5% growth rate of air passenger traffic in 2017 compared to the level of 2016, while in 2018 the traffic growth rate decreased to 6.1%, that is 2.4 percentage points lower compared to the previous year. According to the statistical data published by AAR - the Romanian Airports Association, in Romania air passenger traffic increased by 23% in 2017 compared to 2016, reaching 20,284,022 passengers and by 7.76% in 2018 compared to 2017, reaching the level of 21,858,724 passengers. The passenger growth of Sibiu International Airport (SBZ) is above the national average, in 2017 air passenger traffic increased by 37% compared to the previous year, and in 2018 it increased by 33% compared to 2017, according to published data on the SBZ site.

The World Annual Traffic Forecast 2018 study conducted by ACI International estimates a global average annual growth rate of air passenger traffic of 4.1% by 2040. The AIRBUS company forecasts an average passenger air traffic growth rate of 4.4% per year by 2037 worldwide while BOEING an average annual air traffic growth rate of 4.6% by 2038.

2. Literature review

The theoretical quality of the forecasting models for air traffic of passengers is fundamental for obtaining the most accurate predictions. Existing models for predicting passenger air traffic are divided into several categories:

- **Parametric models (econometric models).** Multiple regression models built on the relationship between passenger traffic and economic, social, demographic variables are intensely used as well as gravity models.

Gravity models' premise is that the passenger is informed and acts rationally from an economic point of view. In this regard, the factors that influence the passenger's behavior to choose a particular airport are: arrival time at the airport, frequency of flights to specific destinations, and flight ticket cost (Brian Graham, 1999).

This category also includes the traffic forecasting methodology used by IATA, which estimates the market air demand based on the socio-economic variables of the market, GDP and adjustment factors: regulations, demand, airlines, competition, substitution (Air Traffic Forecasting Methodology by IATA). It has been concluded that there is generally unidirectional Granger causality from GDP to Revenue Passenger Kilometers (RPK), also used as airline traffic (E. Fernandez, R.R. Pacheco, 2010).

Wenbin Wei and Mark Hansen have used a logarithmic function for estimating the demand of passenger air traffic based on the following independent variables: flight frequencies, number of seats in the aircraft, ticket price, flight distance, number of spokes in the network, airport capacity, airport area population income, number of local passengers who travel from spoke S to hub H by airline A, total number of initiated passenger trips originating from spoke S (Wenbin Wei, Mark Hansen, 2006). A similar model, called the Econometric Dynamic Model (EDM) was used to predict the number of passengers according to economic variables, active population, consumer price index, number of flights, average occupancy rate of hotels, value foreign currency exchange rates values at international arrivals (Rafael Bernardo Carmona-Benítez, Maria Rosa Nieto, Danya Miranda, 2016).

The mixed multinomial logit model (MMNL) has also been used with good results to analyze air travel behavior in airport choice (Stephane Hess, John W. Polak, 2005).

Fuzzy regression analyses is used for estimating passenger air traffic, but also for forecasting cargo air traffic in order to reduce the residual value resulting from the influence of uncertain and unidentified factors in parametric models (T.Y Chou, G.S Liang, T.C Han, 2011; M.S. Liao, G. S. Liang, C. Y. Chen, 2012; V.A Profillidis, 2000).

Similar with the multivariate regression model is the ARIMAX model (Wai Hong Kan Tsui, Hatice Ozer Balli, Andrew Gilbey, Hamish Gow, 2014) which can improve the forecast accuracy by considering the autocorrelation which may exist within the regression residuals.

- **Nonparametric models, namely time series analysis** eliminates the shortcomings of the parametric methods of identifying and predicting all variables that influence the passenger traffic.

Univariate models of time series use only the passenger traffic past data without considering other exogenous variables. These time-sequence models generate better predictive results, according to the study by Ya-Ling Huang and Chin - Tsai Lin who used The Gray Envelope Prediction Model (GEPM) for monthly, seasonal and annual passenger traffic forecasts (Ya-Ling Huang , Chin - Tsai Lin, 2011). The Grey prediction model is suitable for short-term forecasts that do not have a high degree of uncertainty. A Grey-Markov model was applied by Zhang Wei, Zhu Jinfu in 2009 in order to estimate passenger traffic by integrating the Grey model (used for short periods with low uncertainty) in the Markov chain model applicable in dynamic settings.

ARIMA autoregressive model is suitable for short term forecasts and when traffic records regular variations, cyclical seasonality, but generates significant errors when traffic variations are irregular (ACRP, 2007; M. Çuhadar, 2014; A. Danesi, L. Mantecchini, F. Paganelli, 2017).

The SARIMA model is a combination of stochastic seasonal model and ARIMA model and can best describe the fluctuation of passenger flow of the airport terminal which presents a periodic fluctuation. The study carried out by Ziyu Li et al using SARIMA generated predictions of passenger traffic very close to the actual values, the error rate of the model being between 1% and 3% (Ziyu Li, Jun Bi, Zhiyin Li, 2017). Another study developed by Wai Hong Kan Tsui et al for predicting passenger traffic at Hong Kong airport using SARIMA and ARIMAX models has resulted in high accuracy, with very small prediction errors (Wai Hong Kan Tsui, Hatice Ozer Balli, Andrew Gilbey, Hamish Gow, 2014).

For time series characterized by nonlinearity with irregular fluctuations and evolutions, it is recommended to use artificial neural networks (ANNs), support vector machines (SVMs), genetic programming (GP) and ensemble empirical mode decomposition (EEMD) (Y. Bao, T Xiong, Z. Hu, 2012).

The LSTM (long short-term memory) network model, a neural network prediction model is proper for short-term traffic forecasting (Z. Zhao, W. Chen, X. Wu, P.C.Y. Chen, Jingmeng Li, 2017). Neural network forecasting models are also used if there is a nonlinear relationship between the models' variables (T.O. Blinova, 2007).

3. Research methodology

The purpose of the present research is estimating the potential traffic for SIA (Sibiu International Airport, SBZ) which is a regional airport in Romania, for the year 2017. For achieving this, the following SIA internal secondary data was used: (i) monthly number of aircraft departures for every route managed by SIA indifferently of the airline company, and (ii) monthly traffic (departures, arrivals, and transit) of SIA. Secondary data analysis was used as research method, and time series analysis and regression analysis as data analysis techniques. Based on the upper literature review of models (parametric and nonparametric) used for predicting passenger traffic, we propose a mixed-model which uses time-series predictions as input for multiple regression analysis.

In this regard, a two-step process was used in developing the traffic forecasting model: (1) Identifying the proper regression model for traffic estimation based on the number of aircraft departures, and (2) Forecasting the number of aircraft departures for the current routes operated SIA.

(1) Identifying the proper regression model for traffic estimation based on the number of aircraft departures

For estimating the SIA traffic for the year 2017, the causal relationship between the number of aircraft departures and related SIA traffic was taken in consideration. The aircraft departures were divided in internal and external departures, and the related SIA traffic in departure, arrival, and transit traffic.

Simple regression analysis was used as estimation technique; traffic (departure, arrival, and transit) being considered as effect variable, and number of external aircraft departures as causal one. The values of both variables were studied on monthly basis for the period of January 2006 – June 2017. Our decision of including only the number of external aircraft departures in the regression model was because, starting with May 2014

no more internal routes were operated from SIA. The monthly time series of the two regression variables are presented in figures below:

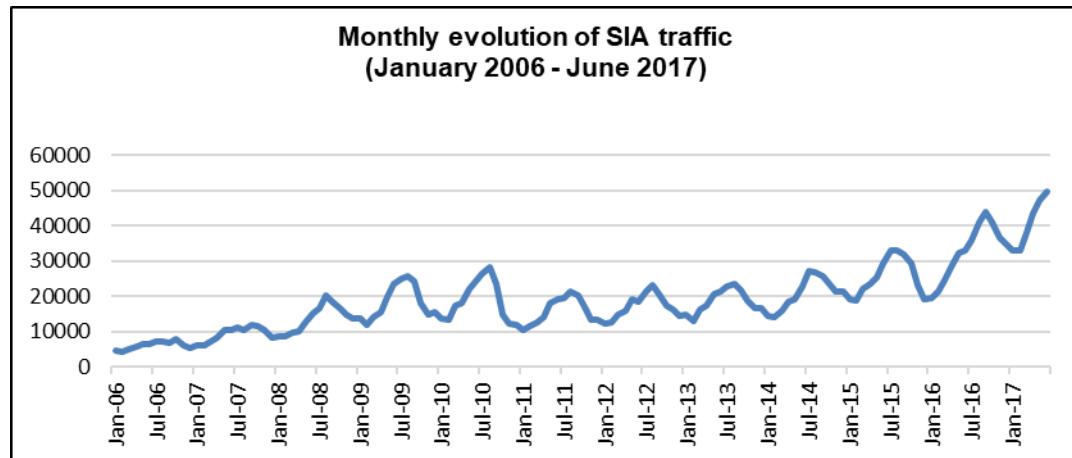


Figure 1 – Monthly evolution of SIA traffic (January 2006 – June 2017)

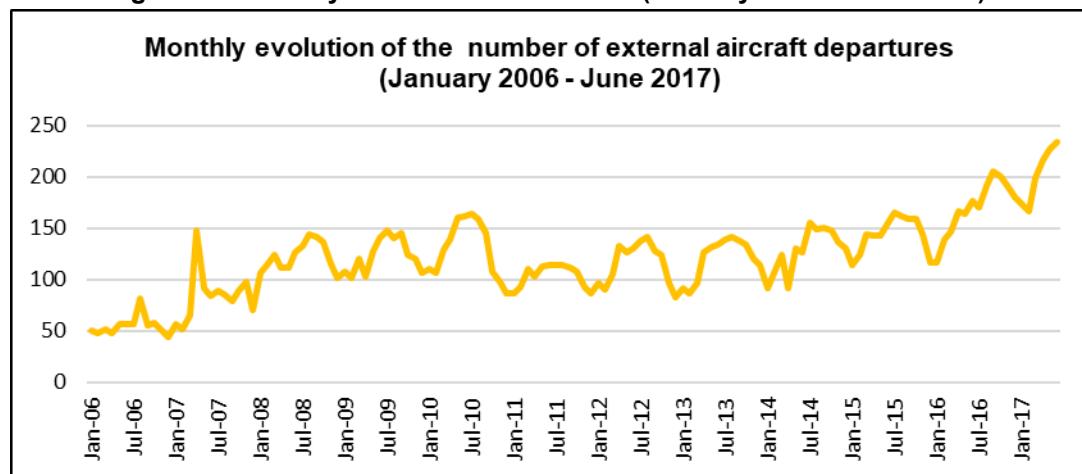


Figure 2 – Monthly evolution of the number of external aircraft departures (January 2006 – June 2017)

Both variables were natural logarithmized (natural-log) for reducing their variance and to linearize them, the initial form of the regression model being presented below:

$$\ln Y_i = \beta_0 + \beta_1 \ln X_i + \varepsilon_i \quad \text{Equation 1}$$

- Y_i = SIA departure traffic for month i ;
- X_i = number of external aircraft departures for month i ;
- ε_i = residual value for month i ;
- i = months considered, 1 being the first month (January, 2006) and 138 being the last one (June, 2017).

After estimating the regression coefficients (valid from a statistical point of view, $\alpha < 0.01$), the observed monthly departure traffic data for 2017 were compared with the estimated ones. Although small, the differences between observed and monthly departure traffic data were not satisfactory for the prediction accuracy. In this regard, a set of simulations were undertaken by including in the initial regression model several Dummy variables linked to some events which have taken place in some periods of time and in other not. The event which has minimized the prediction errors (the upper mentioned differences) was the start of the Sibiu – London (Lutton) route in September, 2015. Based on the above information, the final form of the regression model is:

$$\ln Y_i = \beta_0 + \beta_1 \ln X_i + \beta_2 \text{Dummy}_i + \varepsilon_i \quad \text{Equation 2}$$

- Y_i = SIA departure traffic for month i ;
- X_i = number of external aircraft departures for month i ;
- Dummy_i = 0 if for month i there weren't any aircraft departures from Sibiu to London (Lutton), and 1 if they were;
- ε_i = residual value for month i ;
- i = months considered, 1 being the first month (January, 2006) and 138 being the last one (June, 2017).

	Mean	Std.Deviation	N
$\ln Y_i$	8.89	0.55	138
$\ln X_i$	4.75	0.35	138
Dummy_i	0.16	0.37	138

Table 1 – Descriptive Statistics of the used regression variables

		$\ln Y_i$	$\ln X_i$	Dummy_i
Pearson Correlation	$\ln Y_i$	1.000	0.920	0.595
	$\ln X_i$	0.920	1.000	0.499
	Dummy_i	0.595	0.499	1.000
Sig. (1-tailed)	$\ln Y_i$.000	.000
	$\ln X_i$.000	.000
	Dummy_i			
N	$\ln Y_i$	138	138	138
	$\ln X_i$	138	138	138
	Dummy_i	138	138	138

Table 2 – Pearson correlation values for the used regression variables

The positive, strong and statistically significant correlation ($0.920, \alpha < 0.01$) between the monthly traffic (departure traffic) and the number of external aircraft departures is a hint for the future causality between the two variables. Moreover, a positive, medium and statistical significant correlation is spotted between the monthly traffic (departure traffic) and the Dummy variable ($0.595, \alpha < 0.01$).

R	R ²	Adjusted R ²	Std. Error of the Estimate	Durbin-Watson
0.933	0.871	0.869	0.199	0.904

Table 3 – Overview of the regression model

The high value of the multiple determination coefficient ($R^2=0.871$) underlines the relationship intensity between the effect variable and the two cause variable, meaning that 87.9% of departure traffic variation can be explained by variation of the external aircraft departure number and by the event of introducing the Sibiu - London (Lutton) route. The Durbin-Watson statistics explains the existence or not of autocorrelation of the residual variable's values. The indicators' value should be around 2 for lack of autocorrelation. For our data, the Durbin-Watson statistic is 0.904, roughly one point under the needed threshold. Clearly, there is a hint of positive autocorrelation of the residual values, but the mere existence of it does not influence the regression coefficients, they remain unbiased, although another set of unbiased estimators can be found characterized by a smaller variance.

Model	Sum of Squares	Degrees of Freedom	Mean Square	F	Sig.
Regression	35.912	2	17.956		
Residual	5.333	135	0.040		
Total	41.246	137		454.525	.000

Table 4 – ANOVA table

Analyzing the upper ANOVA table, it can be concluded that the variance of the dependent variable (traffic) is highly influenced be the regression model and lesser by the residual.

Model	Unstandardized Coefficients		t	Sig.	95% Confidence interval for β_i	
	$\hat{\beta}_i$	Std. Error			Beta	Lower Bound
(Constant)	2.677	0.262		10.216	.000	2.159
$\text{Ln}X_i$	1.301	0.056	0.830	23.236	.000	1.190
Dummy _i	0.270	0.053	0.181	5.058	.000	0.164
						0.375

Table 5 – Regression coefficients' values and their statistical significance

Both regression coefficients' values and the constant are statistically significant (Sig. or $\alpha < 0.01$). The Tolerance and VIF Statistics from Table 6 (0.751 and 1.331, respectively) are under 10 (for Tolerance) and above 0.2 (for VIF), not representing therefore some issue regarding the multicollinearity of the independent variables (Field, 2009).

Variables	Correlations			Collinearity Statistics	
	Zero-Order	Partial Correlation	Part Correlation	Tolerance	VIF
$\text{Ln}X_i$	0.932	0.908	0.755	0.751	1.331
Dummy _i	0.556	0.290	0.105	0.751	1.331

Table 6 – Correlations and Collinearity Statistics

Other two regression equation were computed by using the regression model from Equation 2; one for arrival traffic (Y_i = SIA arrival traffic for month i , X_i = number of external aircraft arrivals for month i , and $Dummy_i = 0$ if for month i there weren't any aircraft departures from Sibiu to London (Lutton), and 1 if they were), and the second for transit traffic (Y_i = SIA transit traffic for month i , X_i = number of external, transited aircrafts for month i , and $Dummy_i = 0$ if for month i there weren't any aircraft departures from Sibiu to London (Lutton), and 1 if they were). The statistical tables and interpretation are similar to the departure traffic model.

(2) Forecasting the number of aircraft departures for the current routes of SIA

The second step in predicting SIA traffic for the year 2017 assumed applying another forecasting technique for estimating the number of aircraft departures of the actual routes operated by SIA for the remaining months of 2017 (July till December). On this subject, for every route currently operated by SIA, the monthly number of aircraft departures from the start of route till June 2017 was modelled for estimating the number of aircraft departures for the remaining months of 2017 (July till December). Holt-Winters models (Chatfield, Yar, 1988) were applied for modelling the data series (number of aircraft departures) specific for every route. Thus, based on the data-series' graphical representation of every route, a multiplicative or an additive Holt-Winter model was used to estimate the number of aircraft departures for the remaining months of the year 2017. Every model's parameters for the level equation, seasonality equation, and trend equation were optimized by minimizing the squared differences (errors) of the number of aircraft departures observed and predicted for the first 6 months of 2017.

Based on the presented methodology, equation 2 was used to predict SIA traffic (departure traffic, arrival traffic, and transit traffic) for the last 6 months of 2017 (July till December) using the number of aircraft departures (estimated for the considered months with the Holt-Winters models) and the Dummy variable as independent variables. The results indicated a SIA departure traffic of 260,905 Pax, SIA arrival traffic of 250,165 Pax, and a SIA transit traffic of 35,506 Pax. Summing up the three estimates, our SIA traffic prediction for 2017 is 546,576 Pax. The traffic statistics published by SIA for the year 2017 (<https://www.sibiuairport.ro/en/info/statistici-trafic-aerian/>) indicate an actual total traffic of 533,306 Pax, meaning that our prediction overestimated the real traffic with only 2.4%. Taking out the actual and predicted transit traffic (remaining total actual traffic of 503,906 Pax and predicted of 511,070 Pax), our model overestimates the actual traffic only by 1.42%.

4. Conclusion

Passenger traffic is one of the most important factors which guides the strategic and tactical decisions of both the airport and airline company management. In this regard, accurate estimation of passenger traffic will optimize an airport's financial planning for the periods to come. The present paper describes a prediction model of passenger traffic which was applied for estimating SIA' traffic for the year 2017. A multiple regression model

was considered for estimating the traffic based on two independent variables: (1) daily departures and (2) a Dummy variable. Moreover, Holt-Winters method (additive and multiplicative) was used to estimate the daily departures. The predicted total passenger traffic overestimates the actual total traffic with only 2.4% and the actual total traffic without the transit traffic with only 1.42%.

5. References

- Airport Cooperative Research Program (ACRP), (2007). Airport Aviation Activity Forecasting.
- Antonio Danesi, Luca Mantecchini and Filippo Paganelli, (2017). Long-Term And Short-Term Forecasting Techniques For Regional Airport Planning, *ARPN Journal of Engineering and Applied Sciences*, VOL. 12, NO. 3, FEBRUARY 2017, ISSN 1819-6608.
- Bent Flyvbjerg , Mette K. Skamris Holm & Søren L. Buhl (2007). How (In)accurate Are Demand Forecasts in Public Works Projects?: The Case of Transportation, *Journal of the American Planning Association*, 71:2, 131-146, DOI:10.1080/01944360508976688.
- Brian Graham, (1999). Airport-specific traffic forecasts: a critical perspective, *Journal of Transport Geography* 7 (1999) 285-289.
- Chatfield, C., Yar, M. (1988): Holt-Winters Forecasting: Some Practical Issues, *Journal of the Royal Statistical Society. Series D (The Statistician)*, Vol. 37, No. Special Issue: Statistical Forecasting and Decision-Making, pp. 129 – 140.
- Elton Fernandes, Ricardo Rodrigues Pacheco, (2010). The causal relationship between GDP and domestic air passenger traffic in Brazil, *Transportation Planning and Technology*, 33:7, 569-581.
- Field, A. (2009): Discovering Statistics using SPSS – Third Edition, *Sage Publications Inc.*, London.
- IATA - Air Traffic Forecasting Methodology by IATA, https://www.epd.gov.hk/eia/register/report/eiareport/eia_2232014/html/Appendix%202.1.pdf
- Mao-Sheng Liao , Gin-Shuh Liang , Chin-Yuan Chen, (2012). Fuzzy grey relation method for multiple criteria decision-making problems, *Quality Quantity*, DOI 10.1007/s11135-012-9704-5.
- Murat Çuhadar, (2014). Building Proper Forecast Model for Daily Air Passenger Demand: A Study of Antalya International Airport, *Conference Proceedings International Antalya Hospitality Tourism and Travel Research Conference* held in Antalya from 9-12 December 2014.
- Pavithra Parthasarathi, David Levinson, (2010). Post-construction evaluation of traffic forecast accuracy, *Transport Policy* 17 (2010) 428–443.
- Rafael Bernardo Carmona-Benítez, María Rosa Nieto, Danya Miranda, (2016). An Econometric Dynamic Model to estimate passenger demand for air transport industry, *Transportation Research Procedia* 25 (2017) 17–29.
- Regina R. Clewlowa, Joseph M. Sussman , Hamsa Balakrishnan, (2014). The impact of high-speed rail and low-cost carriers on European air passenger traffic, *Transport Policy* 33(2014)136–143.
- Stephane Hess, John W. Polak, (2005). Mixed logit modelling of airport choice in multi-airport regions, *Journal of Air Transport Management* 11 (2005) 59–68.
- Tsung-Yu Chou, Gin-Shuh Liang , Tzeu-Chen Han, (2013). Application of fuzzy regression on air cargo volume forecast, *Qual Quant* (2013) 47:897–908, DOI 10.1007/s11135-011-9572-4.
- T. O. Blinova, (2007). Analysis of possibility of using neural network to forecast passenger traffic flows in Russia, *Aviation*, 11:1, 28-34.
- V.A. Profillidis, (2000). Econometric and fuzzy models for the forecast of demand in the airport of Rhodes, *Journal of Air Transport Management* 6 (2000) 95-100.

- Wai Hong Kan Tsui, Hatice Ozer Balli, Andrew Gilbey, Hamish Gow, (2014). Forecasting of Hong Kong airport's passenger throughput, *Tourism Management* 42 (2014) 62-76.
- Wenbin Wei, Mark Hansen, (2006). An aggregate demand model for air passenger traffic, in the hub-and-spoke network, *Transportation Research Part A* 40 (2006) 841–851.
- Ya-Ling Huang, Chin-Tsai Lin, (2011). Developing an interval forecasting method to predict undulated demand, *Qual Quant* (2011) 45:513–524, DOI 10.1007/s11135-010-9317-9.
- Yukun Bao, Tao Xiong, and Zhongyi Hu, (2012). Forecasting Air Passenger Traffic by Support Vector Machines with Ensemble Empirical Mode Decomposition and Slope-Based Method, *Hindawi Publishing Corporation, Discrete Dynamics in Nature and Society*, Volume 2012, Article ID 431512, 12 pages, doi:10.1155/2012/431512.
- Ziyu Li, Jun Bi, Zhiyin Li, (2017). Passenger Flow Forecasting Research for Airport Terminal Based on SARIMA Time Series Model, *1st International Global on Renewable Energy and Development (IGRED 2017)*, IOP Conf. Series: Earth and Environmental Science 100 (2017) 012146 doi:10.1088/1755-1315/100/1/012146.
- Zhang Wei, Zhu Jinfu, (2009). Passenger Traffic Forecast based on the Grey-Markov Method, *Proceedings of 2009 IEEE International Conference on Grey Systems and Intelligent Services*, November 10-12, 2009, Nanjing, China.
- Zheng Zhao, Weihai Chen, Xingming Wu, Peter C. Y. Chen, Jingmeng Liu, (2017). LSTM network: a deep learning approach for short-term traffic forecast, *IET Intell. Transp. Syst.*, 2017, Vol. 11 Iss. 2, pp. 68-75, © The Institution of Engineering and Technology 2017.
- ACI statistics, <https://aci.aero/data-centre/airport-statistics-infographics/>
- ACI Europe statistics,
<https://www.aci-europe.org/component/downloads/downloads/5350.html>
- AIRBUS statistics, <https://www.airbus.com/aircraft/market/global-market-forecast.html>
- BOEING statistics,
<https://www.boeing.com/resources/boeingdotcom/commercial/market/commercial-market-outlook/assets/downloads/cmo-2019-report-final.pdf>
- EUROSTAT statistics,
[https://ec.europa.eu/eurostat/statistics-explained/index.php/Air transport statistics#Progressive growth in air transport of passengers in the course of 2017](https://ec.europa.eu/eurostat/statistics-explained/index.php/Air_transport_statistics#Progressive_growth_in_air_transport_of_passengers_in_the_course_of_2017)
- Romanian Airports Association statistics,
<https://www.airportaar.ro/crestere-trafic-de-pasageri-pe-aeroporturile-din-romania-cu-23-pe-anul-2017/>
- Sibiu International Airport Statistics, <https://www.sibiuairport.ro/en/info/statistici-trafic-aerian/>