

# Modeling Demand for Air Cargo in the Colombian Context

Miguel Figueroa Loaiza<sup>1</sup>, Roberto Porto Solano<sup>2</sup>, Ricardo Simancas<sup>2</sup> and Victor Higuera Ojito<sup>2</sup>

<sup>1</sup>MSc (c), Universidad del Norte, Barranquilla - Colombia

<sup>2</sup>Researcher, Corporación Universitaria Americana, Barranquilla - Colombia

**Abstract**— In recent years the growth of air cargo has accelerated, necessitating the assessment of the challenges that this growth will in the future, in order to overcome them and continue contributing to the economic development of the country. Considering the above, this paper proposes several models to estimate demand for air cargo in Colombia, obtained through the use of methodologies such as linear regression and neural networks, which can be used to characterize the current demand and to forecast future demand scenarios given certain contexts to be set. For the estimation of the models presented, information airfreight demand of the top 19 airports in the country was used (in terms of cargo shipped), registered by the Special Administrative Unit of Civil Aeronautics in Colombia (Aerocivil) from 2005 to 2014, besides socioeconomic information on the areas of influence of such airports, obtained mostly by the Administrative Department of National Statistics (DANE) of the same country. Finally a comparison between the results obtained by each modeling methodology, finding better results with neural network models is established.

**Keywords**-cargo; freight modeling; neural networks; linear regression

## I. INTRODUCTION

As expressed in [1], demand for air cargo transport has shown significant growth in recent years, thanks to the strength that has made economic growth, improved political stability and free market generated by open skies policies. Globalization, digitization, aviation and time-based competition make the worlds of commerce and management of supply chains are changing rapidly. High-tech products usually small, lightweight and compact, components and assembled products are shipped increasingly internationally by air quickly and flexibly. Some of the factors that have favored the growth of air cargo worldwide have been the liberalization of air rights, improving the quality of customs, and reducing corruption.

As the Colombian context air transport is one of the most important services in the economy. For one, the country's topography hinders the operation through other transportation, hindering the integration and connectivity of routes and territories. In recent years, growth in Colombia has accelerated, according to [2] Fedesarrollo, among 101 countries, the Colombian air transport held in 2009 ranked 38th in terms of passenger kilometers, and ranked 22nd in tonne-kilometers, while compared with countries in Latin and South American region, Colombia ranked fourth and third respectively.

This entity establishes that from the perspective of aggregate demand intermediate consumption, defined as consumption of air transport made the other sectors of the economy to carry out its activity, representing 35% of total demand, followed by the final consumption made by households, with 27%. The participation of such transport in national GDP remained almost constant at an annual average of 0.5% between 1990 and 2009, which shows that the demand for air transport with the economy evolves.

Therefore, air transport has been gaining greater importance in the development of the country, so to forecast demand for air cargo transport for that type in the following years is essential for future challenges in terms of infrastructure and management. This article based on historical data on air cargo handled in Colombia, obtained from two government agencies: The Special Administrative Unit of Civil Aeronautics (Aerocivil) [3] and the National Administrative Department of Statistics (DANE) [4], linear regression models propose and Neural Networks for estimating air freight, which will serve as a tool for process description and woods, transport such country.

## II. MODELING AIR FREIGHT COLOMBIA AND THE WORLD

According to [1] air freight, trade and GDP have a direct and interdependent relationship, based on the definition of gross domestic product (GDP) as the monetary value of final goods and services produced by an economy in a given period. The authors determined that the demand for air freight will increase as the volume of trade and economic activity increase. They also indicated that foreign direct investment is a good variable to be used in analyzing the growth of air cargo traffic.

In this paper, the authors developed models to forecast the demand for airfreight in future years, in order to determine the required terminals to be prepared for the development of long-term air transport capacity. For this study, the demand for air cargo is analyzed in Taiwan Taoyuan International Airport (TTIA) using system dynamics models. System dynamics is appropriate in the context of the problem is the complexity prediction uncertainty, by modeling and simulation of complex issues and analysis of nonlinear behavior in time.

In [5] analyzed the future demand for air cargo in China and the consequences this will have for the system infrastructure. They used an econometric methodology to determine the relationship between GDP and demand for air cargo. They GDP projections obtained by analyzing trends and

projections of the government of China and recognized institutions.

Most studies conducted in the field consider only the geo-economic characteristics of airports and their hinterlands as explanatory variables in the models; however, in a study conducted by [6] developed a gravity model of the flow of air cargo to try to incorporate more factors that could influence the flows of international air cargo airport, by reference to a base panel data airfreight services on scheduled routes in Taiwan Taoyuan International Airport during the years 2004-2007. In conclusion it was found that the population, the rate of air cargo and three dummy variables, including one related to membership in the regional economic bloc " Chinese Circle "(an informal partnership between Hong Kong, Macao, Taiwan and mainland China), the Open Sky and colonial connections established long ago, agreements are the main determinants of international air cargo flows to and from Taiwan population and income per capita regional included to represent the economic and demographic characteristics of the areas surrounding airports.

In [7] sought to analyze the relationship between economic trends and air cargo traffic in China to provide aggregate of air cargo from China using an econometric method and extrapolate specific projections at major airports in China demand. An econometric approach with a log-log form which was applied linear regression, setting a model for forecasting aggregate air cargo demand from China was used. The model makes projections on the basis of the causal link between air cargo traffic from China and economic activities, the related information is obtained from data time series and economic projections for three different scenarios, pessimistic, most likely And optimistic. The variables included in the model were the GDP, the country's international trade and the flow of foreign direct investment.

Regression analysis is one of the most studied and applied econometric techniques models. For different reasons, first, the linear regression is suitable for modeling a wide variety of relationships between variables, and is used in many applications such as generation passenger load generation, among others, plus they are easy to interpret. It is expected that the linear regression is used to model a linear relationship between a continuous variable and dependent variables and one or more continuous and independent [8].

To predict the flow of international transport of goods through the ports and lands of Jordan, Syria and Lebanon, in [9] developed a model of international freight (IFTM), this model illustrates well the behavior of exporters and importers of goods in international multimodal network. The model is constructed such that exporters of basic products make decisions about where and how to transport their goods; decisions destination, mode, transshipment and routing are made.

Because, the time series of air transport generally has a complex behavior, different techniques have been used for estimation of air transport such as artificial neural networks. For example, it has developed a model for the analysis of air passengers, where it was found that the neural processing goes beyond the traditional econometric approach and offers the

generalization in the behavior of time series, even when not only small samples [10].

Also, we have developed new hybrids, such as those that combine the analysis of singular spectrum, based on the matching network fuzzy inference system and improved particle swarm optimization for predicting air traffic short-term approaches. It aims to identify and extract the trend and seasonality in demand for air transport [11].

### III. METHODOLOGY FOR ESTIMATION OF MODELS

For the realization of this article, a zoning of the country was conducted. Areas used for modeling corresponded to the departments which are the 19 busiest airports worldwide air cargo in the country, this in order to capture trends cargo movement at the national level, considering also that could be obtained historical information on relevant socioeconomic variables that serve to explain the load demand within the model. The lack of this information caused several airports they had to be discarded.

In this sense, the information was used to construct the database was categorized by department and each period. In this study, the information used covers the period 2005-2014 the years and was included in modeling and generated tons of cargo per year.

Most of the information that the database was built for the estimation of the models was provided by the National Administrative Department of Statistics. The variables considered in the study were socioeconomic variables both the department and the people who live there.

One of the important variables in the models was the value tariffs for cargo transportation. These values were taken from the website of a Colombian airline [12]. In this information about the price of one kilogram of mobilizing general cargo between different origin-destination pairs that were available it was found. Importantly, because of this limitation, it was necessary to simplify and assumed that all cargo handled was a general, so a single type of rate was used from each source (assumed average value invariant over time).

In addition, records of cargo moved between origin-destination pairs every month from 2005 to 2014, information that the total annual load generated was determined from the departments included in the study were used. The motion information domestic freight was obtained from the Civil Aviation of the nation.

Once consolidated database, we proceeded to estimating trip generation models. To do so, the methods chosen were linear regression and neural networks.

### IV. LINEAR REGRESSION MODELS

Different models were estimated by evaluating different combinations of socioeconomic variables, production and population of the departments considered in the study. The variables that better behavior presented were the production and economic, as if the department is producer or flowers, gross domestic product per capita agricultural production area,

woods area, among others. Estimated models linear regressions are presented below:

Regression Model 1:

$$\text{Cargo} = \text{Intercept} + \beta_{\text{Rate}} * \text{Rate} + \beta_{(P\_Flowers)} * P\_Flowers + \beta_{(A\_Agricultural)} * A\_Agricultural + \beta_{(A\_Woods)} * A\_Woods + \beta_{GDP} * GDP$$

Regression Model 2:

$$\text{Cargo} = \text{Intercept} + \beta_{\text{Rate}} * \text{Rate} + \beta_{(P\_Flowers)} * P\_Flowers + \beta_{(A\_Woods)} * A\_Woods + \beta_{GDP} * GDP$$

Regression Model 3:

$$\text{Cargo} = \beta_{\text{Rate}} * \text{Rate} + \beta_{(P\_Flowers)} * P\_Flowers + \beta_{(A\_Woods)} * A\_Woods + \beta_{GDP} * GDP$$

The variables used in the regression models are estimated present below:

- **Rate:** Variable representing the cost of sending a ton of cargo from one airport to any destination in the country, an invariant average value over time is assumed. It is expected that within the estimate, this variable is negative, obtaining that higher the rate, the lower the load generation from that department or airport.
- **P\_Flowers:** Binary variable indicating whether production and / or flower trade occurs within the department. This variable is 1 when flower production is given, and 0 but. It is expected that this variable is positive, ie when there is Floreal production load generation is greater.
- **A\_Agricultural:** Variable representing the surface area in hectares intended for agricultural production of a department

intended to be transported by air. It is expected that this variable is negative given that agricultural products are not shipped by air transportation, which increased agricultural production to lower airfreight is generated.

- **A\_Woods:** Variable representing the area in hectares for woodlands and production derived from woods, which will be transported by air inputs. It is expected that the higher the load larger area generated.
- **GDP:** Variable representing the per capita gross domestic product per department. It is expected that this variable is positive given that higher gross domestic product represent greater economic capacity load generation and economic activities.

The models evaluated are shown in the following table:

TABLE I. LINEAR REGRESSION ESTIMATES

Linear regression models to estimate generation air cargo					
Model 1		Model 2		Model 3	
Dependent variable	Cargo (Ton)	Dependent variable	Cargo (Ton)	Dependent variable	Cargo (Ton)
Constant	8036,101 t: 2,834	Constant	7294,481 t: 2,611	Fixed	
Independent Variables					
Rate (\$/Ton)	-0,0051 t: -4,684	Rate (\$/Ton)	-0,0053 t: -4,906	Rate (\$/Ton)	-0,003 t: -4,684
Producer Flowers	6170,368 t: 3,637	Producer Flowers	6134,763 t: 3,606	Producer Flowers	5740,359 t: 3,337
Agricultural area (He)	-0,0146 t: -1,409	Woods area (He)	0,0147 t: 3,225	Woods area (He)	0,017 t: 3,743
Woods area (He)	0,01951 t: 3,434	GDP per capita	0,0006 t: 3,982	GDP per capita	0,00073 t: 5,152
GDP per capita	0,00062 t: 4,095	_____		_____	
Regression model fits					
R^2	0,302	R^2	0,294	R^2	0,416
R^2 Adjusted	0,283	R^2 Adjusted	0,279	R^2 Adjusted	0,402

According to the results of linear regression models estimated above, it can be seen that the linear regression model 1 Although all estimated variables turn out to have the correct sign, the variable A\_Agricultural despite having the correct sign, not it is significant. Regarding the regression model 2, all variables found to be significant and present the correct sign, however this model is the one with the lowest R^2. In the regression model 3 variables included are the same that were considered in the regression model 2, with the difference that the model was set constant given that this turned out not to be

significant. This model also estimated that the variables have the expected sign and be all statistically significant, it is the best fit presents.

## V. ARTIFICIAL NEURONAL NETWORKS

It is an artificial intelligence technology capable of identifying the causal link between the entry and exit through an iterative training, and uses this ratio to obtain predictions [13]. An example of this type of model is presented below:

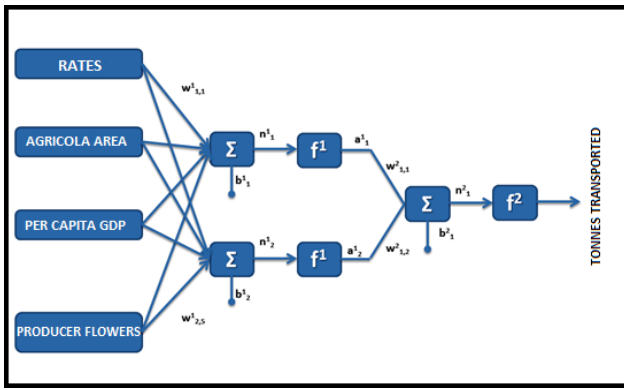


FIGURE I. ARTIFICIAL NEURAL NETWORK EXAMPLE

In this model four input variables and two output one connecting neurons are presented. Moreover, these models show result, different coefficients used in the algorithm to make its prediction. The meaning of each of them is as follows:

- $w^1$  y  $w^2$  weight matrices are in the first and second layer, respectively.
- $b^1$  y  $b^2$  are the vectors of bias in the first and second layer, respectivamente
- $p$  is the input vector.
- $a^1$  y  $a^2$  vectors are output in the first and second layer respectively.
- $n^1$  y  $n^2$  are network entries in the first and second layer respectively.
- $f_1$  y  $f_2$  they are the hidden activation function and output activation function, respectively.

In this study, different models were estimated by evaluating different combinations of socioeconomic, production and population departments variables considered in the study. The variables that were presented better performance and economic production, the same used in the regression models. Estimated models neural networks are presented below:

- Model 1.

$$CA(ton) = f_2 \left\{ [w_2] f_1 \left[ [w_1] \begin{pmatrix} Rate \\ Agricultural\ area \\ Per\ capita\ GDP \\ Producer\ flowers \end{pmatrix} + [b_1] \right] + [b_2] \right\}$$

where CA is the air cargo handled at the airport.

- Model 2.

$$CA(ton) = f_2 \left\{ [w_2] f_1 \left[ [w_1] \begin{pmatrix} Rate \\ Population \\ Per\ capita\ GDP \\ Woods\ area \end{pmatrix} + [b_1] \right] + [b_2] \right\}$$

- Model 3.

$$CA(ton) = f_2 \left\{ [w_2] f_1 \left[ [w_1] \begin{pmatrix} Rate \\ Agricultural\ area \\ Livestock\ area \\ Occupied \\ Rate \end{pmatrix} + [b_1] \right] + [b_2] \right\}$$

For these models used a supervised learning, ie, that the value of the response variable to the set of input variables had. The training algorithm Levenberg-Marquardt. Hidden activation function used was the hyperbolic tangent activation function and output used was linear.

The results of the estimated models are presented:

TABLE II. RESULTS OF NEURAL NETWORK MODELS

	Model 1	Model 2	Model 3
<b>Distribution of the sample</b>			
<b>Training</b>	70	70	60
<b>Validation</b>	15	20	20
<b>Test</b>	15	10	20
<b>Network structure</b>			
<b>No. Neurons</b>	4	4	2
$w^{11}$	0.0007	-0.9224	19.316
	-0.4273	-0.9418	-73.625
	-12.102	22.939	-27.278
	0.9883	23.162	0.3980
	-----	0.1779	26.157
$w^{12}$	-15.549	-47.573	0.1228
	24.942	-51.921	-0.2059
	-0.3991	-25.447	11.346
	-48.034	-24.622	128.833
	-----	-73.733	-18.367
$w^{13}$	-81.290	-0.2451	-----
	-52.404	33.144	
	32.409	33.570	
	41.494	30.243	
	-----	-15.274	
$w^{14}$	31.747	-0.3005	-----
	15.702	19.246	
	-50.915	0.4187	
	-10.571	-0.4559	
	-----	-0.4643	
$b^1$	-28.825	23.431	-25.340
	-31.864	-34.987	88.989
	48.144	0.2638	-----
	-44.421	-10.735	
$w^{21}$	-11.959	0.0461	0.1127
	-0.2346	-0.0435	99.510
	-0.0335	-0.0172	-----
	-0.0441	0.9506	
$b^2$	-19.411	-0.0210	90.590

In addition the results of the model fits in each stage are attached:



Model 1.

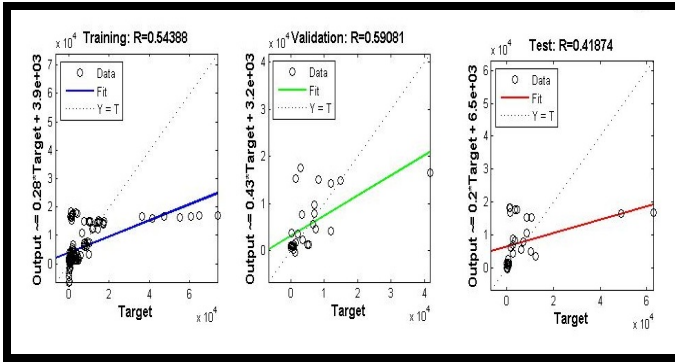


FIGURE II. ADJUSTING STAGES OF MODEL 1

Model 2.

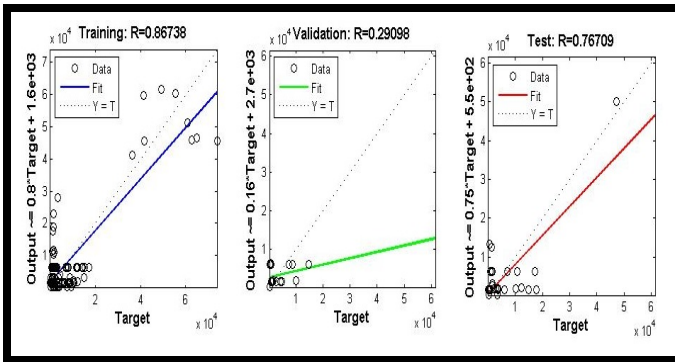


FIGURE III. ADJUSTING STAGES OF MODEL 2

Model 3.

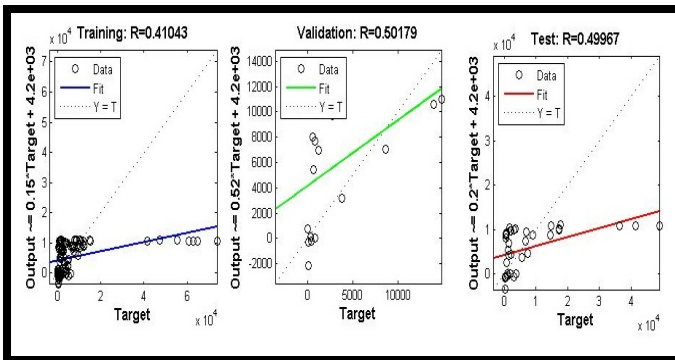


FIGURE IV. ADJUSTING STAGES OF MODEL 3

VI. CONCLUSIONS

In the work presented above models to estimate annual demand for air cargo in the context of the Colombian transportation developed methodologies linear regression and neural networks for estimating them, taking as explanatory variables some of which were found to be relevant in the review of the literature were used. On the other hand variables that influence freight transport in the context of the country they were included, as is for example the production of flowers in some departments.

Within three models presented linear regression variables such as the rate charged for the carriage of cargo, gross domestic product per capita (GDP per capita) and consideration of whether the department (zone of influence) in question is found or not producing flowers have a clear impact on the overall demand for air cargo. It was found that the increase in rates has a negative effect on demand, while GDP growth and flower production has a positive effect, generating an increase in demand for air transport. The results were as expected and consistent with those found in the literature. Although the incidence is not as clear as in the case of the other variables, the area of forests and agricultural production also proved influential.

In the case of neural networks the same variables used in the linear regression in order to verify that even using different estimation model, the influence thereof is clear were used. As expected, the neural network models presented more accurate than those provided by linear regression models, which can be seen when comparing the settings models obtained results. In addition it was found that Neuronal networks models were better than other models when represent the nonlinear behavior of the demand for air cargo in Colombia.

In future estimates could consider some different additional variables, which in this case were not included for lack of information as they are indicators of trade and industry, and manufacturing of specific products, such as the case of flowers, may have an influence.

REFERENCES

- [1] Suryani, E., Chou, S., Chen, C. Dynamic simulation model of air cargo demand forecast and terminal capacity planning. *Simulation Modelling Practice and Theory*, vol. 28, p. 27-41, 2012.
- [2] Olivera, M., Cabrera, P., Bermudez, W., Hernández, A. El impacto del transporte aéreo en la economía colombiana y las políticas públicas, 2011. Recovered on 10/06/2015 de [http://www.repository.fedesarrollo.org.co/bitstream/11445/165/1/CDF\\_No\\_34\\_Abril\\_2011.pdf](http://www.repository.fedesarrollo.org.co/bitstream/11445/165/1/CDF_No_34_Abril_2011.pdf)
- [3] Civil, A. Aeronáutica Civil. Recovered on 09/06/2015, <http://www.aerocivil.gov.co>
- [4] DANE. Departamento Administrativo Nacional de Estadística. Recovered on 09/06/2015, [www.dane.gov.co](http://www.dane.gov.co).
- [5] Jiang H., Ren L., Hansman R.J. Market and infrastructure analysis of future air cargo demand in China. *En AIAA's 3rd Annual Aviation Technology, Integration, and Operations (ATIO) Forum*. p. 6770, 2003.
- [6] Hwang, C., Shiao, G. Analyzing air cargo flows of international routes: an empirical study of Taiwan Taoyuan International Airport. *Journal of Transport Geography*, vol. 19, no 4, p. 738-744, 2011.
- [7] Wu, M., Morrel, P. China's Air Cargo Demand: Future Market Developments and Implications. *Aviation*: p. 1-13, 2010.
- [8] Washington, S. P., Karlaftis, M. G., & Mannering, F. L. *Statistical and econometric methods for transportation analysis* (2da. Edición ed.). Boca Raton, Florida, United States of America: Chapman & Hall/CRC, 2011.
- [9] Serag, M. S.; Al-Tony, F. E. Modeling international freight transport through the ports and lands of Arab countries. *Alexandria Engineering Journal*, vol. 52, no 3, p. 433-445, 2013.
- [10] Alekseev, K. P. G.; Seixas, J. M. A multivariate neural forecasting modeling for air transport—Preprocessed by decomposition: A Brazilian application. *Journal of Air Transport Management*, vol. 15, no 5, p. 212-216, 2009.
- [11] Xiao, Y., Liu, J., Hu, Y., Wang, Y., Lai, K., & Wang, S. (2014). A neuro-fuzzy combination model based on singular spectrum analysis for air

transport demand forecasting. *Journal of Air Transport Management*, vol. 39, p. 1-11, 2014...

- [12] LAN. LAN CARGO. Recovered on 09/06/2015, <http://www.lancargo.com/es/ruta-colombia/tarifas>
- [13] M. Yasin C & A. Tortum. Modelling car ownership in Turkey using neural networks. En *Proceedings of the Institution of Civil Engineers-Transport*. Thomas Telford Ltd, p. 97-106, 2009.