International Journal of Aviation, Aeronautics, and Aerospace

Volume 9 | Issue 3 Article 7

2022

Air passenger demand forecast through the use of Artificial Neural Network algorithms

Juan Gerardo Muros Anguita
Universidad de Granada, gerardo.muros@gmail.com
Oscar Díaz Olariaga
Universidad Santo Tomas - Colombia, diaz_olariaga@yahoo.es

Follow this and additional works at: https://commons.erau.edu/ijaaa

Part of the Other Aerospace Engineering Commons, and the Transportation Engineering Commons

Scholarly Commons Citation

Muros Anguita, J. G., & Díaz Olariaga, O. (2022). Air passenger demand forecast through the use of Artificial Neural Network algorithms. *International Journal of Aviation, Aeronautics, and Aerospace, 9*(3). https://doi.org/10.15394/ijaaa.2022.1726

This Article is brought to you for free and open access by the Journals at Scholarly Commons. It has been accepted for inclusion in International Journal of Aviation, Aeronautics, and Aerospace by an authorized administrator of Scholarly Commons. For more information, please contact commons@erau.edu.

Plans for developing different components of an airport system depend to a great extent on the levels of activity that are anticipated. In order to plan facilities and infrastructures of an airport or system/group of airports, and to be able to satisfy future needs, it is essential to predict the level and distribution of demand in the various components of the airport system (Transportation Research Board [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 make air traffic estimates as a preliminary step to planning and designing airport facilities, be it an airport, an airport system, or a network (Horonjeff et al., 2010; Wells & Young, 2004).

Understanding future demand patterns enables the airport planner to evaluate the future performance of the airport and, thereby, recommend consistent development programs, so that the costs associated with these development plans are estimated and the sources and level of income to support future capital investments are projected (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, thus understanding the entire airport network of a region or country (International Civil Aviation Organization [ICAO], 1987; Janic, 2021, 2009).

To assess the characteristics of future demand, it is necessary to develop reliable predictions of airport activity. There are many factors that will affect demand, so planners preparing demand forecasts or updating existing forecasts should consider, in addition to historical aeronautical data (air traffic), local socio-economic data (historical series) such as national wealth, purchasing power of the inhabitants, demographics (population), industrial production, consumer price index, exchange rate (of the local currency against the US dollar), etc. These indicators have a great influence on the behavior of air traffic demand (García Cruzado, 2013; Horonjeff et al., 2010; ICAO, 2006; Rodríguez et al., 2020).

The objective of this research is to make a short and medium-term forecast of air passenger demand. For this, Colombia (complete network of airports as a whole) has been used as a case study, with the particularity that the analysis includes data on demand for 2020, which has been severely affected by the COVID-19 pandemic. This analysis will serve to estimate as a complementary result (but of great interest) an approximate date of recovery of both the volume of demand and its growth trend to the pre-pandemic period (2019). To achieve this objective, and as a calculation tool, a model derived from Artificial Neural Networks of the ConvLSTM2D type (<Conv> of Convolutional and <LSTM2D> of long-short-term memory) is developed. This type of architecture is a hybrid between Convolutional Neural Networks (CNN), very useful for the extraction of invariant patterns in their spatial position, and Recurrent Neural Networks (RNN), very appropriate for the extraction of patterns within their temporal context, as it is the case of time series forecast (Hermans & Schrauwen, 2013; Malhotra et al., 2015; Millstein, 2018; Sewak et al., 2018; Yang et al., 2015). These prediction techniques, based on Machine Learning/Deep Learning (ML/DL), can incorporate more elements of analysis and, therefore, potentially be more efficient (Ketkar & Moolayil, 2021). Other advantages of ConvLSTM2D networks, compared to classical methods based on autoregression, is that they admit a multivariate treatment (several input features) (Calin, 2020; Pedrycz & Chen, 2020). Finally, when applying this type of neural model (ML/DL ConvLSTM2D) to the case study of the present research, it comes to represent the abstract knowledge model, inferred from the learning of historical patterns in the time series of air traffic, which predicts the future evolution of these time series.

Literature Review

In the scientific literature there are various approaches based on ANN for forecasting passenger demand and air travel. The first is a hybrid neural model with data preprocessing by decomposition of variables, which allows improving the performance of the network and, thus, optimizing the results in the forecast of demand (Alekseev & Seixas, 2009). The second is the Ensemble Empirical Mode Decomposition (EEMD) based on Support Machines Vector (SMV), with a modeling framework that incorporates slope-based methods to constrain the problem of the final effect that occurs during the process of change of the EEMD; in other words, an EEMD-Slope-SVMs model (Bao et al., 2012).

The third is the use of Backpropagation Neural Network (BNN) to improve the accuracy of the demand forecast. (Chen et al., 2012; Kuo & Chen, 2010). The fourth is a hybrid approach VMD-ARMA/KELM-KELM, which consists of the Variational Mode Decomposition (VMD) in an Autoregressive Moving Average (ARMA) model and in a Kernel Extreme Learning Machine (KELM). This means that the VMD is first adopted to break down the original data into various functions, in order to reduce its complexity. Then ARMA and KELM models are used to forecast the stationary and non-stationary components, respectively, and the final result is integrated by another KELM model, which incorporates the forecast results of all components (Jin et al., 2020).

The fifth approach consists of the combined use of Machine Learning and Support Vector Regression (SVR) (Plakandaras et al., 2019; Sulistyowati et al., 2018). The sixth is the Nonlinear Vector Autoregression Neural Network (NVARNN) approach, based on MIV. This approach consists of using: (1) a Mean Impact Value (MIV) method from a neural network to identify and extract input variables, and (2) NVARNN to deal with the irregularity and volatility of the time series (Sun et al., 2019). The seventh is the ensemble forecast modeling based on Singular Spectrum Analysis (SSA), this implies that the original time series is decomposed into three components: trend, seasonal oscillations and irregular component. The trend is predicted using a Generalized Regression Neural Network

(GRNN), while seasonal oscillations are predicted using Radial Basis Function Network (RBF Network) (Xiao et al., 2015).

The eighth approach consists of using a Multilayer Perceptron Architecture (MLP) with a feedback propagation algorithm (Dingari et al., 2019). The ninth entails using evolutionary metaheuristic algorithms (Mostafaeipour et al., 2018); the tenth, uses time-delayed feedback neural networks (Blinova, 2007), and, finally, eleventh uses recurrent neural networks (RNN) based on long-short-term memory (LSTM) (Gupta et al., 2019).

Data and Methodology

In the methodology, prior to the development of the model, the data (or dataset) related to the annual historical time series of each of the six independent variables (or selected features) were prepared. These are the aeronautical variables (national/domestic passenger or international passenger), and the socioeconomic variables (at the national level) usually used in the calculation of forecasts in air transport (Díaz Olariaga and Girón, 2020; Rodríguez et al., 2020), which are: GDP (Gross Domestic Product), population, IPI (Industrial Production Index), IPC (Consumer Price Index), and TRM (US Dollar-Colombian Peso official exchange rate). The data of this research correspond to the country-case study selected, Colombia; historical data cover a continuous period of 42 years (1979-2020); aeronautical data were obtained from the statistical system of the Colombian Aeronautical Authority (Aerocivil, 2022) and socioeconomic data were obtained from related official sources existing in the country (BRC, 2022; DANE, 2022).

Although the data –both air traffic and socioeconomic— for 2020 are very low compared to previous years due to the reasons explained previously, they are included in the calculation with the aim of imposing the model to learn the COVID-19 effect patterns.

Finally, as far as data is concerned, an important part of preparing the dataset has consisted in extracting the temporal sequences of input to the model and split up them into subsequences relative to the number of previous steps (*lags* or past) of each independent variable or feature, necessary to make predictions (*outs* or future steps). Although the code trains the complete historical time series of 42 years, the data are prepared sequentially into temporal subdivisions of the steps of the previous years (called *n_lags*) before considering future prediction steps (called *n_outs*). These *n_lags* grouping can be seen as time context to extract time cycles features. In addition, these values have been hyper-parametrized in the algorithm to study the cost-sensitivity to them. Likewise, output data corresponding to national or international passenger are extracted from the dataset, with the historical subsequences corresponding to the number of steps in future years to be predicted. For simplicity, the same number of steps from previous years has been considered as future steps to be predicted (*n_lags* = *n_outs*).

The model that is developed in the present research, which is one derived from artificial neural networks, is called ConvLSTM2D (<Conv> of Convolutional and <LSTM2D> of long-short-term memory) (Hermans & Schrauwen, 2013; Malhotra et al., 2015; Millstein, 2018; Sewak et al., 2018; Yang et al., 2015). What the Conv2LSTM2D model aims, as expressed in equation (1), is that once the model has been trained or adjusted in the connections (*weights*) within the neurons of the different layers, symbolized by the function $f_{\text{model forecast}}$, this predicts (output variables denoted $\widehat{\boldsymbol{Y}}$) the *n_outs* temporary outputs corresponding to the *i* features ($\hat{y}_{it+n\_outs}$), from the input of the *n_lags* observations of previous times of the *i* features ($x_{it-n_{lags}}$).

$$\hat{y}_{it+n\_outs}, \hat{y}_{it+n\_outs-1}, \dots, \hat{y}_{it} = f_{model\ forecast}\left(x_{it-1}, x_{it-2}, \dots, x_{it-n_{lags}} \right) \quad (1)$$

To describe the equations of the ConvLSTM2D model, we begin by outlining the basic architecture of ANN, on which the ConvLSTM2D model is based, called Multilayer Perceptron Networks (MLP) (Vang-Mata, 2020). These are made up of several layers of neurons, with a first layer corresponding to the input (associated with the **X** tensor) and a last output layer (associated with the **Y** tensor), the hidden layers are interspersed between them. Each layer contains a variable number of neurons (or nodes) that can be activated with activation functions of various types (Hastie et al., 2013).

The ConvLSTM2D model adopted here is of the Deep Learning (DL) type and has an encoder-decoder structure in which the **X** inputs are first encoded, with hybrid layers of convolutional (CNN) and recurrent (RNN) characteristics (Aggarwal, 2018; Bianchi et al., 2017; Blokdyk, 2017; Mandic & Chambers, 2001). On the other hand, the decoding is performed only with the recurring layers of long-short-term memory (LSTM). The coding provided by the convolutional layers facilitates the representation of the invariances of the sequences of the features of the inputs, while the contribution of the LSTM cells in the coding allows memorizing the context of the temporality of the inputs. The decoding part of the model is performed only by the LSTM layers, which allows extracting and identifying the patterns of the features and the present temporality, such as trend and seasonality (Pedrycz & Chen, 2020). Finally, the output of the LSTM networks is delivered to other dense layers of neurons, MLP type (where all neurons are connected between them) which are the ones that conduct the learning of the feature patterns extracted by the previous encoder-decoder block, until reaching the last layer, where the output tensor **Y** is obtained. In summary, the application of encoding-decoding with convolutional CNN and recurring LSTM layers in the ConvLSTM2D model is equivalent to producing double deep learning (Brownlee, 2018; Trifa et al., 2017; Yang et al., 2015).

The group of equations (2) that define the layers of the ConvLSTM2D model, according to the notation represented in Figures 3, 4, and 5, were proposed as an improvement of the traditional RNN (Donahue, 2015; Hu et al., 2019), which managed the hidden states of the cell to map the output from the input. LSTM networks were proposed to solve the problems that appeared by gradient vanishing and exploding (Hochreiter & Schmidhuber, 1997). The superiority of the ConvLSTM2D layered architecture lies in its ability to handle short and long-term memory provided by the LSTM architecture by managing the three types of gates: input, forget and output, plus the new gate coming from the contribution of the convolutional layers with which they are combined to produce the new ConvLSTM2D layer (Shi et al., 2015). Then:

$$i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + \; W_{ci} \circ c_{t-1} + b_i) \qquad (2a)$$

$$f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \qquad (2b)$$

$$o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ c_t b_o) \qquad (2c)$$

$$g_t = tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \qquad (2d)$$

$$c_t = f_t \circ \; c_{t-1} + i_t \circ \; g_t \qquad (2e)$$

$$h_t = O_t \circ \tanh(c_t) \qquad (2f);$$

in which: i_t, f_t, o_t represent respectively the input gate, the forget gate and the output gate of the ConvLSTM2D cell; x_t, the temporary t inputs of the cell; h_{t-1} and c_{t-1}, the output and the state respectively of the previous cell; * is the convolution operator and *W* is the convolution filter; with $W_{\bullet i}$ and b_i the equivalent of the weights and biases of the input gate; $W_{\bullet f}$ and b_f , the equivalents to the weights and biases of the forget gate; $W_{\bullet o}$ and b_o, those at the output gate, and $W_{\bullet c}$ and b_c the weights and biases of the cell state. The dimensions and processing of the ConvLSTM2D layers, while being analogous to those of the LSTM cell, are different, since the hybrid ConvLSTM2D cell also performs the convolution operation (Hu et al., 2019). Finally, the operator $\circ$ denotes the Hadamard product and σ the activation function.

It is worth mentioning that, in the algorithm developed for the present research, the activation function σ chosen has been the one of the Rectified Linear Unit Function (ReLU) (Nair & Hinton, 2010), because it offers better results for the selected case study than the function of the hyperbolic tangent (*tanh*) present in the equations described in (2).

Before training the model, the compilation method is implemented. In compilation method the model's arguments are defined, such as: gradient optimizers, cost or error function, and metric evaluation. The next step in the methodology involves training the ConvLSTM2D model. Training the ConvLSTM2D model is under supervised learning, that involves defining the error metric, which measures the ‘distance’ or norm between the value of the real outputs of the time series used from the **Y** tensor, and the predicted or estimated $\widehat{\mathbf{Y}}$ value by the ConvLSTM2D model, after applying it to the tensor input **X** (Goodfellow et al., 2016). In this research, the mean square error (MSE) has been used for the cost function. Equation (3) defines this error with the norm 2 || ||, where $\boldsymbol{E}$ is the error and $\boldsymbol{n}$ is the number of samples from the mean.

$$E = \frac{1}{n}\sum_{x}\|(y(x) - \hat{y}(x))\|^2 \qquad (3)$$

In ML, Backpropagation (BP) is the fundamental process for the automatic adjustment of the model weights going backward from the cost function (Goodfellow et al., 2016). BP calculates the gradient of said total error (in this case the MSE) compared to each of the weights or connections of the neurons of each layer of the model. For this, the chain rule in the derivation applies, in order to propagate said error backwards, to distribute the error among all the weights of the model, neuron by neuron, layer by layer. With this, the new adjustment of weights of the model is sought to produce a smaller error when it is iterated with the next batch of input samples **X**. If the solution converges, said error becomes smaller or asymptotic, as is the case analyzed here. The group of equations (4) shows the derivation chain rule in ANN networks (Aggarwal, 2018; Hassoun, 1995):

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial w_{ij}} = \frac{\partial E}{\partial o_j}\frac{\partial o_j}{\partial net_j}\frac{\partial net_j}{\partial w_{ij}} \qquad (4a)$$

$$\frac{\partial net_j}{\partial w_{ij}} = \frac{\partial}{\partial w_{ij}}\left(\sum_{k=1}^{n} w_{kj} o_k\right) = \frac{\partial}{\partial w_{ij}}\, w_{ij} o_i = \, o_i \qquad (4b)$$

The problem of minimum searching of the cost function are implemented through different optimizers, generically known as Stochastic Gradient Descent (SGD) (Kingma & Ba, 2014). These optimizers try to avoid that the solution is trapped in some of the local minima that the cost function presents when it is not a convex function. To this end, each of them has implemented different strategies to seek the global optimum.

For the present research, one of the most efficient optimizers has been used, given its high convergence speed and robust stability, known as Adaptive Moment Estimation (ADAM) (Kingma & Ba, 2014). This is an improved version of the RMSProp that implements a variable/adaptive speed (or learning step) and a moment or inertia that tries to avoid getting stuck in the local minima of the cost

DOI: https://doi.org/10.15394/ijaaa.2022.1726

function. The group of equations (5) describe the update of the weights by the ADAM optimizer (Kingma & Ba, 2014):

$$m_w{}^{(t+1)} \leftarrow \beta_1 m_w + (1+\beta_1)\, \nabla_w\, l^{(t)} \tag{5a}$$

$$v_w{}^{(t+1)} \leftarrow \beta_2 v_w{}^{(t)} + (1-\beta_2)\, (\nabla_w\, l^{(t)})^2 \tag{5b}$$

$$\widehat{m}_w = \frac{m_w{}^{(t+1)}}{1+\beta_1{}^{(t+1)}} \tag{5c}$$

$$\hat{v}_w = \frac{v_w{}^{(t+1)}}{1+\beta_2{}^{(t+1)}} \tag{5d}$$

$$w^{(t+1)} \leftarrow w^t - \eta\, \frac{\widehat{m}_w}{\sqrt{\hat{v}_w}+\epsilon} \tag{5e};$$

in which $w^{(t+1)}$ represents the weights in iteration t+1; η, the learning speed or step; $\nabla_w\, l^{(t)}$ represents the gradient of the cost function l with respect to the weights; ϵ, a minimum value to avoid dividing by 0 (in Keras 10^{-7} is used); β_1 and β_2 are the forget factors for the gradient or the first and second moments, respectively; $m_w{}^{(t+1)}$ is the moving mean of the weights or first moments in the iteration $t+1$, and $v_w{}^{(t+1)}$ is the moving variance (or second moment) of the weights in the iteration $t+1$. It should be noted that in the notation of equations (5) the superscripts with the variable *t* correspond to the iterations (or learning cycles), while the notation of the subscripts *t* of the rest of the equations corresponds to the different instants of time dataset features from the time series.

In this work, both the number of iterations and the learning batch size have been hyper-parametrized. A batch is a unit or subset of the dataset from which the error of the cost function is estimated and, therefore, the weights are adjusted. Also, the gradient has the stochastic attribute or name, since it cannot be calculated for the full dataset, due to time, process, and memory computational constraint, then SGD choose stochastic batch instead. All of this makes the learning process sensitive to the batch size, so this value was hyper-parameterized in an algorithm (Calin, 2020; Pedrycz & Chen, 2020; Yi & Tan, 2004).

The next step in the methodology is the evaluation of the model, in which the goodness of the training or adjustment of the model weights is compared and evaluated. The sensitivity of training variations is taken into account trough performing different hyperparameters configurations defining both the model and the training process itself. In the field of ML/DL, the search for the best solution is a heuristic problem, in which different solutions of architectures and model training processes are proposed and tested, extracted from the infinite space of possible solutions to, then, measure their performances and be able to select among them the best models and configurations (Aggarwal, 2018; Patterson & Gibson, 2017).

For this reason, in the evaluation process it is required, on the one hand, to define the space of the hyperparameters used to configure the model and its training, and, on the other hand, to define the metrics used to evaluate and compare the accuracy of the results. As evaluation metrics for regression problems the Mean Absolute Percentage Error (MAPE), which measures the prediction error as a percentage, has been implemented here. This is an advantage since it provides an intuitive way to evaluate the error of the model; the smaller the MAPE, the better the prognosis (Kim & Kim, 2016; Ren & Glasure, 2009; Rodríguez et al., 2020). Additionally, the Root Mean Square Error (RMSE) was also implemented. To implement this, once the model has been trained with all the data, so that the weights of all the connections of the neurons/layers of the model are finally fixed, the full output values $\mathbf{Y}$ of the dataset are again compared to the predicted or estimated $\widehat{\mathbf{Y}}$ values, after applying the model to the full input values $\mathbf{X}$.

Regarding the ConvLSTM2D model, the hyperparameters initially managed to define the possible space of solutions have been the number of filters per convolutional-recurrent layer, the size of the sliding windows of the convolution, the number of times the encoder of ConvLSTM2D layers is repeated, the presence or absence of specific layers such as dropouts and batch normalization (which help to avoid overfitting and be trained robustly) and, finally, varying the number of neurons in the LSTM layer (Donahue et al., 2015). With all this, a robust model has finally been obtained regarding its performance.

In relationship to the process of training and splitting up the time series dataset into sequences of *n_lags* length in order to predict future sequences *n_outs*, after performing several internal tests the framework analysis has focused on two scenarios: the first one, 6 years (short term) and the second, 12 years (medium term), with the two hyperparameters n-lags and n_outs equal to each other. Given that implementing Keras callbacks techniques, such as EarlyStopping and ModelCheckpoint, related to stopping learning when the threshold error is reached and saving the optimal learning model found in training respectively, the algorithm is, to some extent, independent of the number of iterations used. Finally, the hyperparameters that have been taken into account have been the number of subsequences, the number of features used as inputs to predict national passengers or international passengers, together with previous steps *n_lags* and future steps to predict *n_outs*.

Finally, and as the last stage of the methodology, the prediction of the model is developed, which consists of making the predictions of the future series (*n_outs*) with the trained model, from the series of the previous years (*n_lags*) chosen and from the features taken into account as independent input variables.

For this, applying said prediction to the historical (real) series **Y** and, thus, be able to observe the deviation with the estimated outputs $\hat{\mathbf{Y}}$ or predicted by the ConvLSTM2D model was considered convenient. This prediction process is ultimately the final objective of this methodology, since the development of the algorithm to predict the demand for air passenger for future years is what is sought. Although the objective has been focused on two of the features, national (or domestic) passenger and international passenger (annually), the model could predict for the coming years any of the other five additional features included in the historical data set (as are socioeconomic data).

It should be noted that the model weights for each of the possible combinations between any number of the seven possible output features, depending on the seven input features, is different, since the training $\mathbf{Y} = f(\mathbf{X})$, or adjustment of weights of the ConvLSTM2D model, is done only with the appropriate features of the **X** and **Y** tensors extracted from the historical data set.

Results

Next, the results of the prediction of demand (in the short and medium term) of national (or domestic) and international passengers in the Colombian airport system (in set) obtained with the defined, trained and evaluated algorithm are presented, interpreted, and discussed.

The two hyperparameters to which the possible prediction scenarios have been reduced are those relative to the number of previous input and prediction years, respectively *n_lags* and *n_outs*. On the other hand, said pre-post times are subdivided into two or three subsequences, due to the fact that results are sensitive to those number of subsequences, among other hyperparameter, since they define the temporality context from which the patterns of the short-term cycles may exist within those longer historical periods.

This sensitivity is common on the analysis of ConvLSTM2D networks and also is responsible of the fact that finally the input data set **X** has five dimensions: for the number of samples, for the number of subsequences, for time series, for the number of years per subsequence associated to *n_lags*, and for the number of features selected as independent variables of input tensor **X** (e.g. GDP, Population, IPI, IPC, TRM). On the contrary, output tensor **Y** has 3 dimensions: for the number of samples, for the number of forecast years (*n_outs*) and for the number of dependent variables to be forecasted (domestic or international passenger). In the scenarios considered the time series itself of the dependent variable to be predicted is always taken also as independent input variable to provide greater robustness and predictability to the model.

Before describing the hyperparameters and results achieved in each of the most relevant scenarios, Figure 1 shows the learning curve (adjustment of weights) of the model, relative to the iterations of learning cycles (where a robustness of convergence towards asymptotic values of MAPE is observed, with small fluctuations). While Figure 2 shows the prediction of the historical series with the model already trained where we can realize how the decrease in passenger traffic in 2020 (full pandemic situation) was not predicted, as a result of pre-pandemic values. Both curves have been taken as examples of scenario 1, which is described later.

Figure 1
Evolution of MAPE According to the Learning Curve of the Trained Model Described in Scenario 1. Source: Authors.

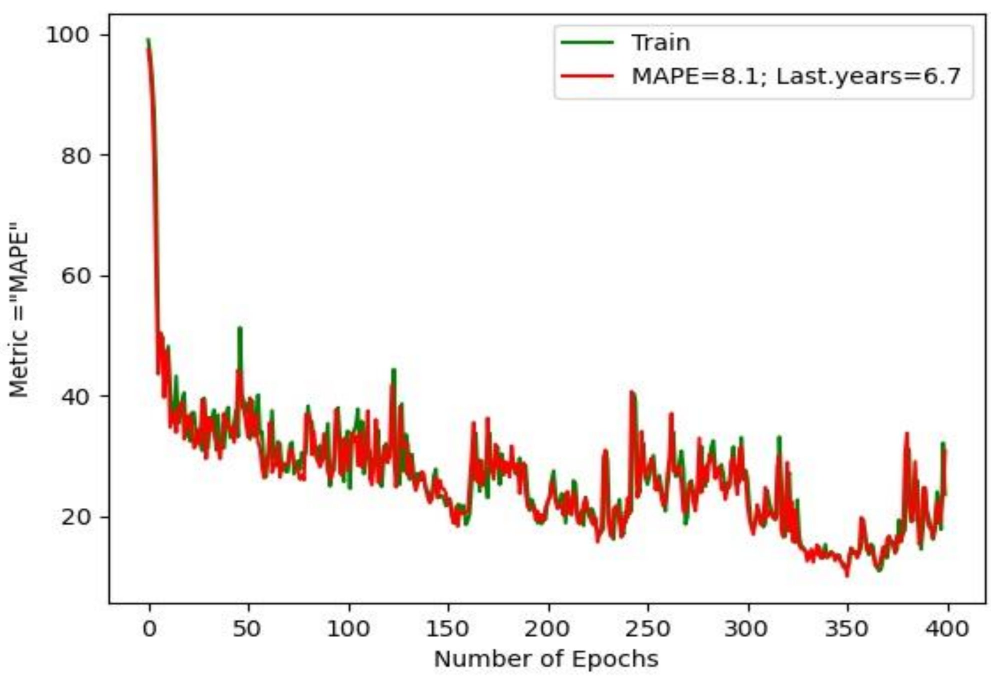
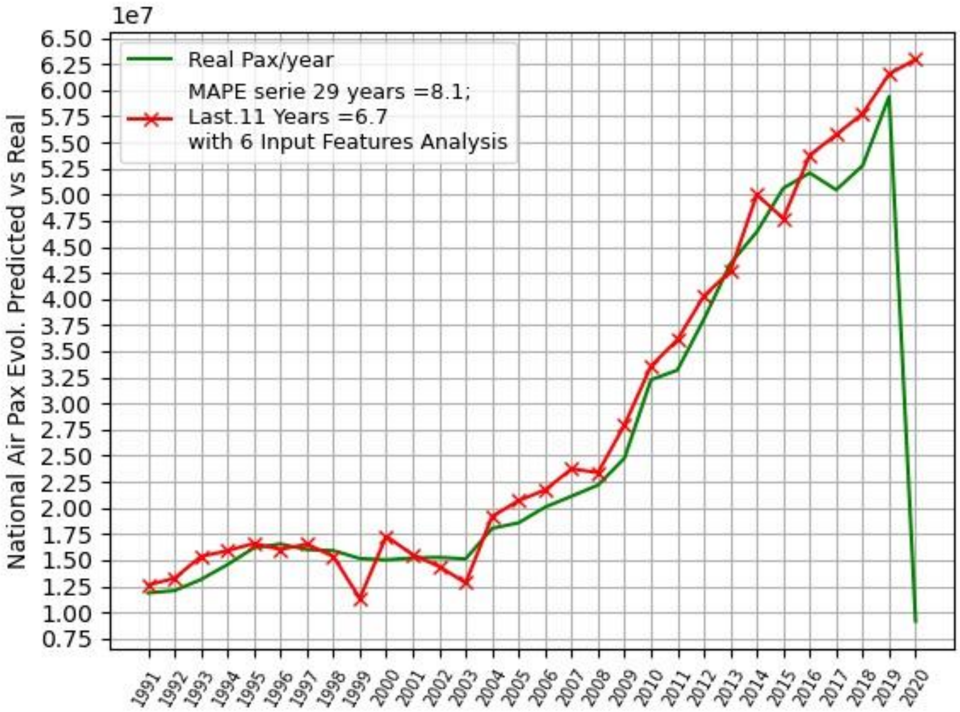


Figure 2
Evolution of Passenger Prediction vs. the Real Historical Series, Started n_lags Years After 1979 of the Model Already Trained and Described in Scenario 1. Source: Authors.



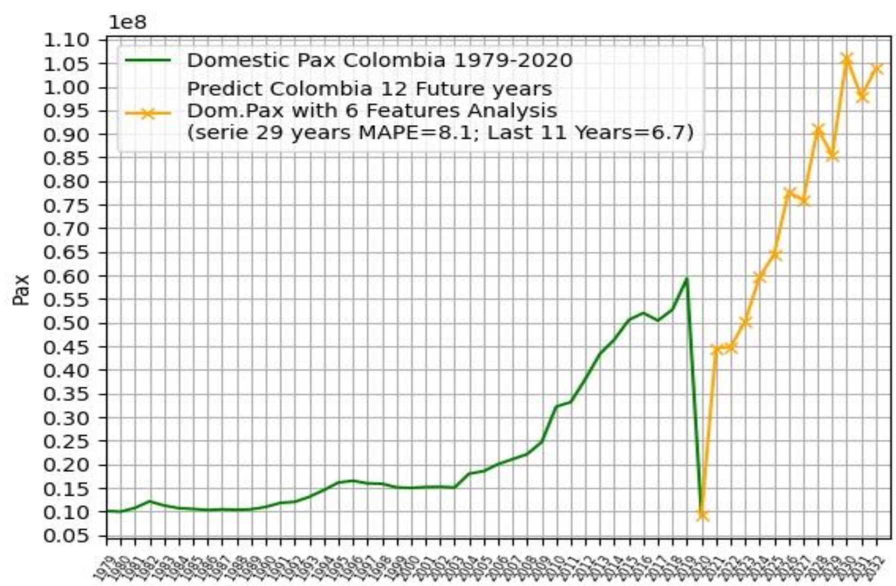
Results are generated and presented in various scenarios in order to deal with the fact that ML/DL models are stochastics in nature. It means results depend on (or are) sensitive to many factors external ones, such as dataset quality (veracity on historical values collected without noise and biases), or the number of inputs (features) and output chosen, but also to the internal ones of ConvLSTM2D model such as the hyperparameters deeply described here as the number of subsequences (2 or 3), length of years considered as previous inputs (*n_lag*) and predicted outputs (*n_outs*), number of model training epochs (or iterations), training batch size, training and test dataset size, optimizers and cost function chosen, number and type of layers implemented and number of neurons per layer, etc.

Scenario 1

A first prediction is presented, based on sequences of the previous 12 years and the following 12 years, from the total historical series, with the time series of six features chosen as independent variables (national passenger, GDP, population, IPI, IPC and TRM) and with national passengers as the only output depending variable to be predicted. The previous 12 years *n_lags* are grouped into three subsequences (or possible sub-cycles) of 4 years each. Figure 9 shows this medium-term prediction with a MAPE value of 6.7% for the last 11 years and 8.1% for the historical series of 29 years (12 years after 1979 are counted as previous steps *n_lags* to predict the 12 years following *n_outs*). In Figure 3 it can be observed that by the beginning of 2024 the traffic (or demand) of national passenger, at the country level, would recover the existing level of the pre-pandemic period (2019), with a recovery of the trend in demand slightly higher than the one of the year 2019, but with oscillations that could indicate potential echoes of the impact of COVID-19.

Figure 3

Forecast of National (or Domestic) Passenger Demand at the Country Level (Colombia), Period 2021-2032; Scenario 1. Source: Authors.

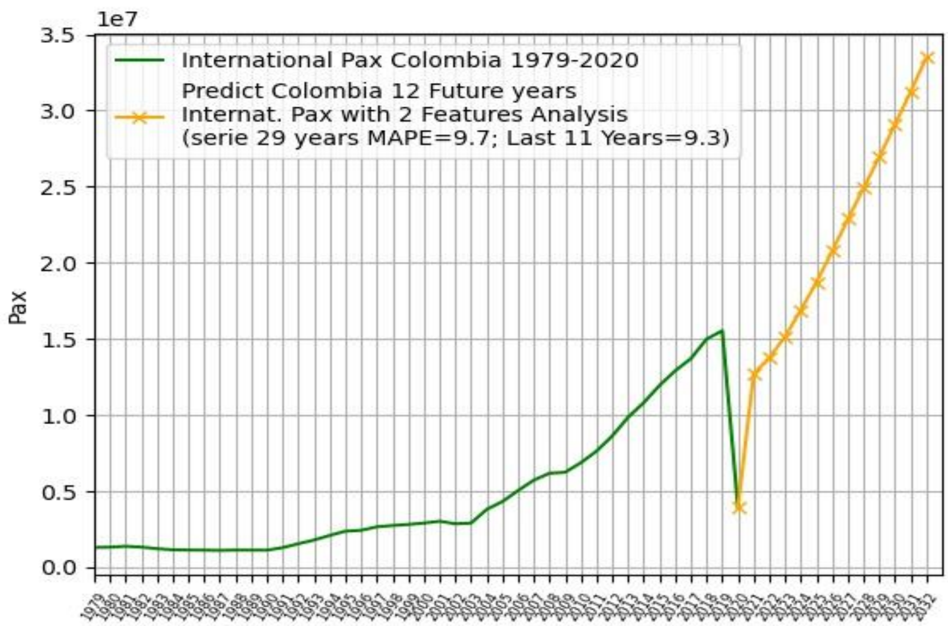


Scenario 2

A second prediction is presented considering sequences of the previous 12 years and the following 12 years, from the total historical series, with the time series of only two features chosen as independent variables (international passenger, GDP) and with international passenger as the only output dependent variable to predict. The previous 12 years *n_lags* are grouped into three subsequences (or possible sub-cycles) of 4 years each. Figure 4 shows this medium-term prediction with a MAPE value of 9.3% for the last 11 years and 9.7% for the historical series of 29 years (12 years after 1979 are counted as previous steps *n_lags* to predict the 12 years following *n_outs*). In Figure 10 it can be observed that by the beginning of the year 2024 the traffic (or demand) of international passenger in Colombia would recover to the level existing in the pre-pandemic period (2019) and with a recovery of the trend in demand very similar to that of the year 2019 and without fluctuations.

Figure 4

Forecast of the Demand of International Passenger in Colombia, Period 2021-2032; Scenario 2. Source: Authors.



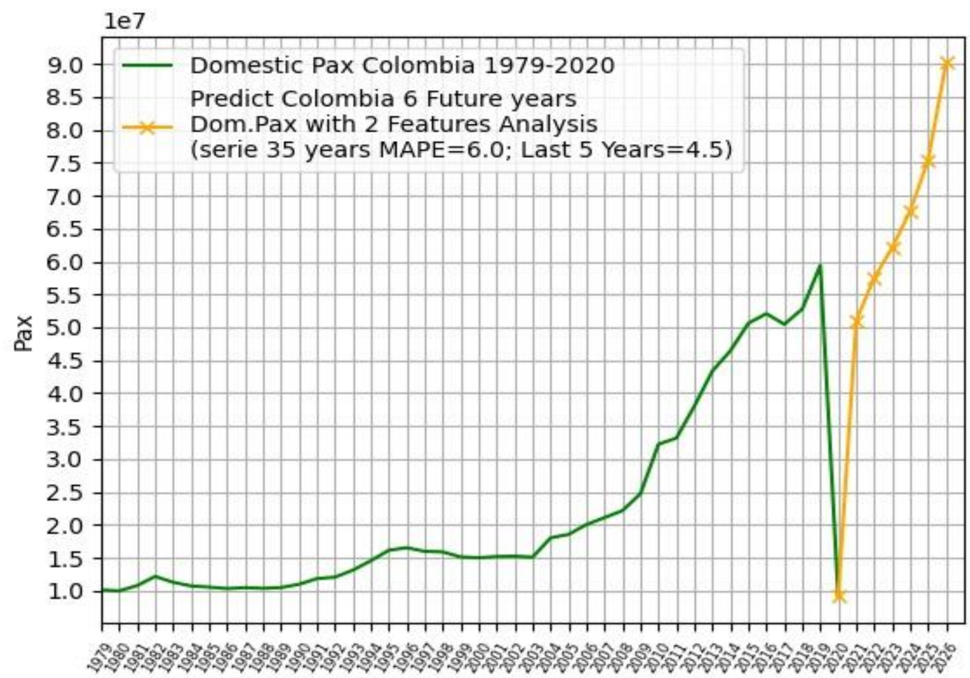
Scenario 3

A third prediction is presented based on sequences of six previous years and the six following years, out of the total historical series, with the time series of only two features chosen independent variables (domestic passenger, GDP) and domestic passenger as the only output dependent variable to predict. The previous six years *n_lags* are grouped into two subsequences (or possible sub-cycles), therefore, of three years each. Figure 5 represents this short-term prediction with a MAPE value of 4.5% for the last five years (2020 is excluded) and 6.0% for the 35-year historical series (starting six years after 1979, the initial year of the available history, to predict the following *n_outs* years). The results show that by the end of

2022 or the beginning of 2023, the traffic (or volume of demand) of national passenger, at the country level, would have recovered the pre-pandemic level (2019).

Figure 5

Forecast of the Demand of National (or Domestic) Passenger at the Country Level (Colombia), Period 2021-2026; Scenario 3. Source: Authors.

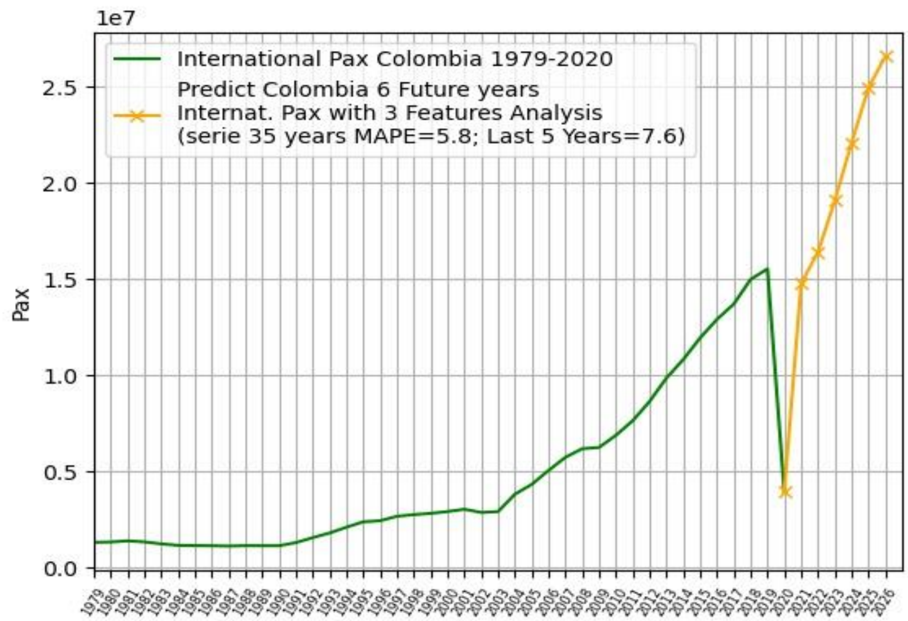


Scenario 4

A fourth prediction is presented, calculated with the trends of the time series of national and international passengers used as input features. To analyze its impact the number of features used as inputs is left at three. Sequences of the previous six years and the following six years have been considered from the total historical series, with the time series of only 3 features chosen as independent variables (international passenger, domestic passenger, GDP) and with international passenger as the unique dependent variable to predict. The previous six years *n_lags* are also grouped into two subsequences (or possible sub-cycles), therefore, of three years each. The results of this calculation are shown in Figure 6. The resulting MAPE values are 7.6% for the last five years (2020 is excluded) and 5.8% for the 35-year historical series (starting six years after 1979, the initial year of the available data history, to predict the following *n_outs* years). The results indicates that by mid-2022 the demand of international passenger, at the country level, would recover the volume of pre-pandemic demand (year 2019) and even with a higher trend (in demand growth).

Figure 6

Forecast of the Demand of International Passenger in Colombia, Period 2021-2026; Scenario 4. Source: Authors.

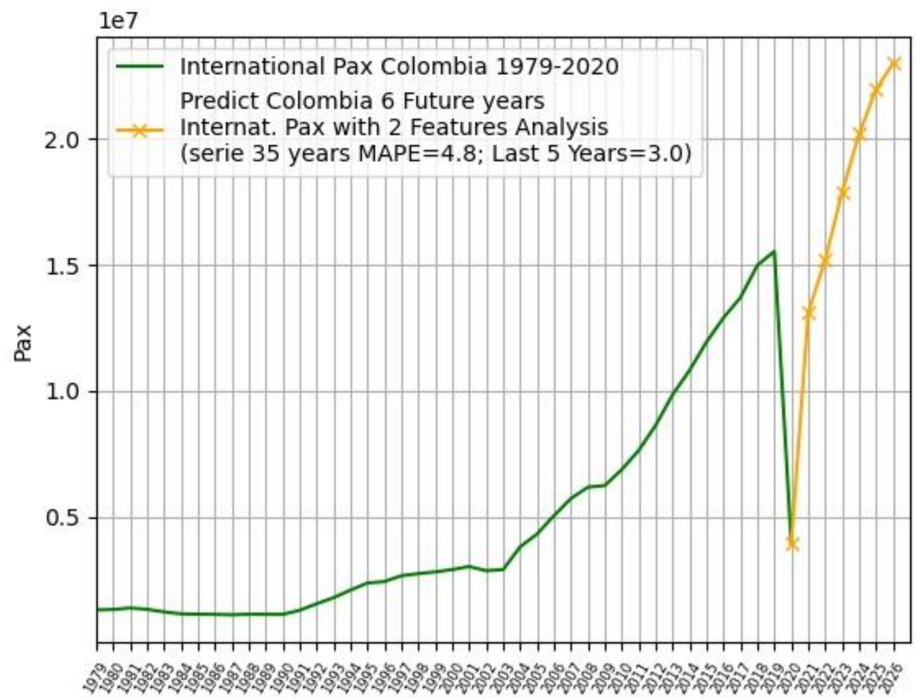


Scenario 5

Finally, a fifth prediction is presented with sequences of six previous years and the six following years, out of the total historical series, with the time series of only two features chosen as a dependent variable (international passenger, GDP) and with international passenger as the only output variable to predict. The previous six years *n_lags* were also grouped into two subsequences (or possible sub-cycles), therefore, of three years each. The results are shown in Figure 7, the MAPE values are 3% for the last five years (2020 is excluded) and 4.8% for the 35-year historical series (starting six years after 1979, the initial year of the available data history, and to predict the following *n_outs* years). In this scenario, the results indicate that by mid-2022, Colombia's international passenger demand would have recovered to the level of pre-pandemic demand (2019).

Figure 7

Forecast of the Demand of International Passengers in Colombia, Period 2021-2026; Scenario 5. Source: Authors.



As a conclusion of the analysis of the results, three observations are made. In the first place, the algorithm developed here has a basic-medium complexity, characterized by the presence of 1,920,009 neural connections, which must be adjusted as unknowns in each iteration during the learning process. This value it is specified for scenario 5 (with two features and six years of forecast) but grow to 3,149,833 neural connections when six features and 12 years of forecast of scenario 1 are handled. Second, as has been exhaustively shown in the description of the ML/DL model of ConvLSTM2D, these are stochastic processes, implicitly originated in the same process of calculating the Stochastic Gradient Descent (SGD), which has been improved here chosen ADAM as loss (or cost function) optimizer (Kingma & Ba, 2014) and, therefore, subject to uncertainties or errors. Third, to constraint the uncertainties associated with the stochastic processes the algorithm has been executed several times, with the same hyperparameters –that process is known as cross-validation techniques (K-fold cross-validation) (Chang & Lin, 2011; Pedrycz & Chen, 2020).

Finally, 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 at the end of 2022 (optimistic scenario). And the results also coincide with the IATA study (2020) which foresees the recovery of worldwide air passenger traffic at the 2019 level by the end of 2022 for domestic passenger and the beginning of 2024 for international passenger.

Conclusions

Firstly, estimating the demand for air passengers at the national level in the short and medium term provides valuable information so that the aviation/air transport planners of a country can well in advance design, plan and implement: (a) development strategies (infrastructures, facilities, equipment, technological modernization, training of technical personnel, etc.); (b) a capital investment calendar (to address the proposed development programs); and (c) related public policies (to consolidate and reinforce the development of the local air transport industry).

Secondly, academics, analysts, planners and decision-makers in the international civil aviation industry around the world are currently developing forecast studies that allow them to know when and how air traffic will recover (mainly the volume of the air passenger demand) to pre-pandemic levels (year 2019), due to the importance of the aviation industry, not only in terms of local/regional/global connectivity, but also because of its contribution to the global economy. The present work had the objective of contributing from the academic point of view to these two mentioned points. On the other hand, this work provides a novel and original geographic character, as one of the first studies (of an academic nature) in the Latin American subcontinent on air traffic forecasting for the post-COVID-19 era.

Regarding the approach, it was decided to develop a model based on an artificial neural network architecture compared to other possible classical statistical techniques, considering that ML/DL prediction techniques can not only incorporate more elements of analysis but also extract more complex patterns from historical time series dataset, without any previous feature engineering and, therefore, be more potentially effective and accurate. In particular, for the analysis of the time series, techniques of encoder-decoder networks of the ConvLSTM2D type have been applied. The ConvLSTM2D model developed here admits a multivariate treatment (the same does not happen with the classic methods based on auto regression). On the other hand, the ConvLSTM2D model developed is of the supervised learning type, which means that the model is trained with tensor inputs and tensor outputs extracted from the historical series of the dataset. The possibility of hybridization of the convolutional CNN networks with the recurrent networks of long-short-term memory LSTM, has allowed the implementation of the

ConvLSTM2D model in the present research (novel in terms of its application for air traffic prediction time series). The application of encoder-decoder blocks with hybrid convolutional-recurrent networks represents a novelty for the treatment of time series. Finally, the results of the application of the model based on ML/DL present very acceptable MAPE values (in the order of 3% to 9%, depending on the scenario), which makes the model developed here a feasible alternative to develop reliable air traffic forecasts, at least in the short and medium term.

References

Aerocivil (2022). *Estadísticas aeronáuticas*. https://www.aerocivil.gov.co/atencion/estadisticas-de-las-actividades-aeronauticas/boletines-operacionales

Aggarwal, C. (2018). *Neural networks and deep learning*. Cham: Springer International.

Alekseev, K. P. G., & Seixas, J. M. (2009). A multivariate neural forecasting modeling for air transport – Preprocessed by decomposition: A Brazilian application. *Journal of Air Transport Management*, 15, 212–216.

Ashford, N., Mumayiz, S., & Wright, P. (2011). *Airport engineering*. John Wiley & Sons.

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*, ID 431512, 1-12. doi:10.1155/2012/431512

BCR – Banco de la República de Colombia (2022). *Estadísticas*. https://www.banrep.gov.co/es/-estadisticas

Bianchi, F., Maiorino, E., Kampffmeyer, M., Rizzi, A., & Jenssen, R. (2017). *Recurrent Neural Networks for Short-Term Load Forecasting*. Cham: Springer.

Blinova, T.O. (2007). Analysis of possibility of using neural network to forecast passenger traffic flows in Russia. *Aviation*, *11*(1), 28-34.

Blokdyk, G. (2017). *Recurrent neural network*. Brendale (Australia): The Art of Service.

Brownlee, J. (2018). *Deep learning for time series forecasting: Predict the future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery, E-book: Jason Brownlee.

Calin, O. (2020). *Deep learning architectures. A mathematical approach*. Cham: Springer.

Chang, C., & Lin, C. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 1-27.

Chen, S., Kuo, S., Chang, K., & Wang, Y. (2012). Improving the forecasting accuracy of air passenger and air cargo demand: the application of back-propagation neural networks. *Transportation Planning and Technology*, *35*(3), 373-392. http://dx.doi.org/10.1080/03081060.2012.673272

DANE - Departamento Administrativo Nacional de Estadística (Colombia) (2022). https://www.dane.gov.co/

Díaz Olariaga, O., & Girón, E. (2020). Análisis de la influencia de la privatización de aeropuertos en el pronóstico de la demanda de pasajeros. El caso de Colombia. *Revista Transporte y Territorio*, 22, 94-113. doi:10.34096/rtt.i22.8408

Dingari, M., Reddy, M., & Sumalatha, V. (2019). Air traffic forecasting using artificial neural networks. *International Journal of Scientific & Technology Research*, 8(10), 556-559.

Donahue, J., Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Darrel, T., & Saenko, K. (2015). Long-term recurrent convolutional networks for visual recognition and description. *IEEE Conference on Computer Vision and Pattern Recognition*, 7-12 July 2015, Boston (MA).

García Cruzado, M. (2013). *Aeropuertos. Planificación, Diseño y Medio Ambiente*. Madrid: Ibergarceta Publicaciones.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

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.

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

Hassoun, M. (1995). *Fundamentals of artificial neural networks*. MIT Press.

Hastie, T., Tibshirani, R., & Friedman. J. (2013). *The elements of statistical learning. Data mining, inference, and prediction*. Springer.

Hermans, M., & Schrauwen, B. (2013). Training and analyzing deep recurrent neural networks. *Working Paper*, Ghent University.

Hochreiter, S., & Schmidhuber, J. (1997) Long short-term memory. *Neural Computation*, 9(8), 1735–1780.

Horonjeff, R., McKelvey, F., Sproule, W., & Young, S. (2010). *Planning and Design of Airports*. McGraw-Hill.

Hu, W., Li, H., Pan, L., Li, W., Tao, R., & Du, Q. (2019). Feature extraction and classification based on spatial-spectral ConvLSTM neural network for hyperspectral images. arXiv:1905.03577 [cs.CV].

IATA. (2020). *Air passenger forecasts: Potential paths for recovery into the medium and long run*. Geneva: Author.

International Civil Aviation Organization. (2006). *Manual on air traffic forecasting. Doc 8991*. Montreal: ICAO.

International Civil Aviation Organization. (1987). *Master planning. Part 1. Doc 9184*. Montreal: ICAO.

Janic, M. (2021). *System analysis and modelling in air transport*. CRC Press.

Janic, M. (2009). *Airport analysis, planning and design: demand, capacity and congestion*. Nova Science Publishers.

Jin, F., Li, Y., Sun, S., & Li, H. (2020). Forecasting air passenger demand with a new hybrid ensemble approach. *Journal of Air Transport Management*, 83, 101744. https://doi.org/10.1016/j.jairtraman.2019.101744

Kingma, D., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv:1412.6980 [cs.LG].

Ketkar, N., & Moolayil, J. (2021). *Deep learning with Python*. Springer.

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, *32*(3), 669-679.

Kuo, S., & Chen, S. (2010). Air passenger and air cargo demand forecasting: applying artificial neural networks to evaluating input variables. *12th WCTR*, July 11-15, 2010. Lisbon (Portugal).

Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2015). Long-short term memory networks for anomaly detection in time series. *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. 22-24 April 2014, Bruges.

Mandic, D., & Chambers, J. (2001). *Recurrent neural networks for prediction: Learning algorithms, architectures and stability*. Wiley.

Millstein, F. (2018). *Convolutional neural networks in Python*. Scotts Valley: CreateSpace Independent Publishing Platform.

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), 5461-5484. https://doi.org/10.1007/s11227-018-2452-0

Nair, V., & Hinton, G. (2010). Rectified linear units improve restricted Boltzmann machines. *27th International Conference on Machine Learning*. June 21-24, 2010, Haifa (Israel).

Patterson, J., & Gibson, A. (2017). *Deep learning: A practitioner's approach*. Sebastopol (CA): O’Reilly Media Inc.

Pedrycz, W., & Chen, S. (2020). *Deep learning: Concepts and architectures*. Cham: Springer.

Plakandaras, V., Papadimitriou, T., & Gogas, P. (2019). Forecasting transportation demand for the U.S. market. *Transportation Research Part A*, 126, 195–214. https://doi.org/10.1016/j.tra.2019.06.008

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), 409.

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

Sewak, M., Karim, R., & Pujari, P. (2018). *Practical convolutional neural networks*. Packt Publishing.

Shi, X., Chen, Z., Wang, H., Yeung, D., Wong, W., & Woo, W. (2015). *Convolutional LSTM network: A machine learning approach for precipitation nowcasting*. arXiv:1506.04214 [cs.CV].

Sulistyowati, R., Suhartono, Kuswanto, H., Setiawan, & Astuti, E. (2018). Hybrid forecasting model to predict air passenger and cargo in Indonesia. *International Conference on Information and Communications Technology (ICOIACT)*, 6-7 March 2018. Yogyakarta, Indonesia.

Sun, S., Lu, H., Tsui, K., & Wang, S. (2019). Nonlinear vector auto-regression neural network for forecasting air passenger flow. *Journal of Air Transport Management*, 78, 54–62. https://doi.org/10.1016/j.jairtraman.2019.04.005

Transportation Research Board. (2002). *Aviation demand forecast: Survey of methodologies*. Washington DC: Author.

Trifa, A., Sbai, A., & Chaari, W. (2017). Enhancing assessment of personalized multi-agent system through convlstm. *Procedia Computer Science*, 112, 249–259.

Vang-Mata, R. (2020). *Multilayer perceptrons: Theory and applications*. Nova Science Pub. Inc.

Wells, A., & Young, S. (2004). *Airport planning & management*. McGraw-Hill.

Xiao, Y., Liu, J., Xiao, J., Hu, Y., Bu, H., & Wang, S. (2015). Application of multiscale analysis-based intelligent ensemble modeling on airport traffic forecast. *Transportation Letters: International Journal of Transportation Research*, 7(2), 73-79. doi: 10.1179/1942787514Y.0000000035

Yang, J., Nguyen, M., San, P., Li, X., & Krishnaswamy, S. (2015). Deep convolutional neural networks on multichannel time series for human activity recognition. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, 3995–4001.

Yi, Z., & Tan. K. (2004). *Convergence analysis of recurrent neural networks*. Dordrecht: Springer.