MODELLING AUSTRALIA'S OUTBOUND PASSENGER AIR TRAVEL DEMAND USING AN ARTIFICIAL NEURAL NETWORK APPROACH

Panarat Srisaeng[1], Glenn Baxter[2]

[1,2] *School of Tourism and Hospitality Management, Suan Dusit University, Huahin Campus, Huahin, Prachup Khiri Khan, Thailand, 77110*

Abstract: This paper focuses on predicting Australia's outbound international airline passenger demand using an artificial neural network (ANN) modelling method. The modelling in the study was based on annual data for the period 1993 to 2016. The model was developed using the input parameters of world GDP, world population growth, world jet fuel prices, world air fares (proxy for air travel cost), Australia's tourism attractiveness, outbound flights, Australia's unemployment levels, the Australian and United States foreign exchange rate and three dummy variables (Sydney Olympics, 9/11 and the 2006 Commonwealth Games). The artificial neural network (ANN) used multi-layer perceptron (MLP) architecture that compromised a multi-layer feed-forward network and the sigmoid and linear functions were used as activation functions with the feed forward-back propagation algorithm. The ANN was applied during training, testing and validation and had 8 inputs, 1 neurons in the hidden layers and 1 neuron in the output layer. The data was randomly divided into three data sets; training, testing and model validation. The best-fit model was selected according to four goodness-of-fit measures: mean absolute error (MAE), mean square error (MSE), root mean square errors (RMSE), AND mean absolute percentage errors (MAPE). The highest R-value obtained from the ANN model is 0.99733, demonstrating that the ANN provided a high predictive capability.

Keywords: air transport, artificial neural network (ANN), Australia, forecasting methods, outbound passengers.

1. Introduction

Because of Australia's relatively remote geographical location, air transport plays a significant role in facilitating the country's commerce and tourism attractiveness. Air and ocean transport are the only two modes available for passengers wishing to travel to/or from Australia. In 2016, sixty-one international scheduled airlines operated services to/from Australia, this included 5 dedicated all-cargo airlines but excludes airlines operating only through code-share arrangements. There was a total of 18,703,064 outbound passengers from Australia in 2016 (Bureau of Infrastructure, Transport and Regional Economics, 2017). The key air travel markets from Australia include China, Japan, New Zealand and the USA.

The airlines serving Australia follow two distinct business models; the full-service network carrier (FSNC), for example, Cathay Pacific and Qantas Airways, and the low-cost carrier (LCC), for example, AirAsia-X and

---

[1] Corresponding author: panarat_sri@dusit.ac.th

Jetstar Airways. Irrespective of the business model that is defined and implemented, forecasting and modelling future air travel demand is regarded as one of the most critical airline management functions (Srisaeng *et al.,* 2015a). This is because future passenger air traffic forecasts influence fleet planning, route network development, investor efficiency, as well as helping the airline to reduce its risk profile through an objective evaluation of the demand side of its business (BaFail *et al.,* 2000).

Forecasting future air travel demand therefore plays an important role in the decision-making and planning for airlines (Srisaeng *et al.,* 2015a). Historically, multiple linear regressions models have been extensively used to model airline passenger demand (Abed *et al.,* 2001; Aderamo, 2010; Bhadra, 2003, Kopsch, 2012; Sivrikaya and Tunç, 2013). In recent years, the use of artificial neural networks (ANNs) has grown rapidly due to their ability to map any linear or non-linear function. ANNs also do not have any associated data assumption requirements (Kunt *et al.,* 2011; Santos *et al.,* 2014). ANNs are also specifically designed to handle a very large, complex systems that have many inter-related parameters (Kalogirou, 2014).

Despite their reported benefits, there have only been a small number of studies undertaken that have used ANNs to model air travel demand. Alekseev and Seixas (2002, 2009) developed ANN-based forecasting models to predict the annual Brazilian air transport passenger demand. In another study, (Chen *et al.,* 2012) employed a back-propagation artificial neural network (BPN) to improve the forecasting accuracy of air passenger and

air cargo demand from Japan to Taiwan. More recently, Srisaeng *et al.,* (2015a, 2015b) proposed and tested ANNs for predicting Australia's domestic and low-cost carrier (LCC) passenger demand, respectively. Despite the significance of Australia's outbound airline services to the country's economy, there has been, to the best of the authors' knowledge, no any reported study that has developed and modelled Australia's outbound air passenger demand using an ANN-approach. Thus, the primary aim of this study is to address this apparent research gap.

The remainder of the paper is structured as follows: section 2 presents the ANN architecture, data collection and variable selection, goodness-of-fit measures and the training and testing of the ANN. Section 4 summarises the empirical results. The study's conclusions and suggestions for future research are presented in section 5.

2. Artificial Neural Network Modelling

2.1. Artificial Neural Network Architecture

Artificial neural networks (ANNs) are a method of using computer software to learn to recognise patterns in given data (Luxton, 2016). ANNs work by creating connections between the mathematical processing elements, which are referred to as *neurons* (Skias, 2006). ANNs capture the inherent information from a considered set of variables and learn from the existing data, even when noise is present (Garrido *et al.,* 2014). Thus, the system learns via a process of determining the number of neurons or nodes and through the adjustment of the weights for the connections based upon the

training data (Skias, 2006). No formulation or a priori model is required (Merkus and Meesters, 2014). A neural network can be trained to perform a particular function by adjusting the values of connections (weights) between the elements (Kunt *et al.,* 2011). During the training process, the ANN can detect complex relationships between the input and output data and perform synthesis (Sivanandam *et al.,* 2006; Sineglazov *et al.,* 2013). Once the ANN has been trained on the sample of the given data-set, it can make estimations through the detection of similar patterns in future data. ANNs can detect similarities in inputs, despite a particular input not ever being seen previously. This property provides ANNs with excellent interpolation capabilities, particularly when the input data may not be exact, that is, noisy (BaFail, 2004).

The most general form of an artificial neural network (ANN) used in forecasting is shown in the following equation:
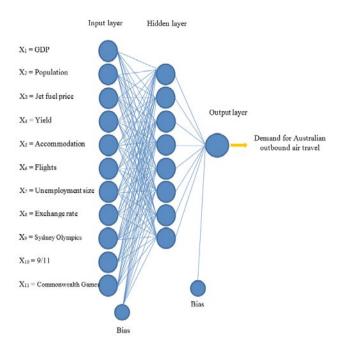
$$Y = F\left[H1\,(x),\, H2\,(x)\, ..., \, Hn(x)\right] + u \quad (1)$$

Where, Y is a dependent variable, X is a set of explanatory variables, F and H's are network functions, and u is a model error term (BaFail, 2004, p. 103).

The artificial neural network model is characterized by a network of three layers: input, output and hidden layers (Sen and Sas, 2014; Yang *et al.,* 2016). Artificial neural networks consist of a large number of simple processing elements called neurons which are organized into several layers and interconnected with each other through synaptic weights. Synaptic weights represent the intensity of the interaction between

every pair of neurons, and the activation functions calculate the potential of every neuron (Garrido *et al.,* 2014; Tiryaki and Aydın, 2014).

The most widely used ANN type for prediction is the Multi-Layer Perceptron (MLP) model (Kale *et al.,* 2016; Rao, 2011; Tiryaki and Aydın, 2014). The MLP is a supervised neural network based on the original simple perceptron model. The MLP is comprised of a system of simple inter-connecting elements, called neurons, cells or nodes. Each of the various inputs included in the model are multiplied by a connection weight. These products are subsequently summed, fed through a transfer function in the ANN to generate a result, and then the model's output (Shirgure and Rajput, 2014, p. 134).

Figure 1 presents the study's 3-layer back propagation network (Remennikov and Mendis, 2016; Shirgure and Rajput, 2014). The first layer is the input layer and corresponds to the problem input variables with one node for each input variable. The second layer is the hidden layer used to capture non-linear relationships among variables. The third layer is the output layer used to provide predicted values (Dulikravich and Colaço, 2015). The number of neurons in the input layer is equal to the number of input variables or independent variables, and the number of output neurons is equal to the number of output variable(s) or dependent variable(s). The input layer receives the initial values of the variables, the output layer shows the results of the network for the input, and the hidden layer carries out the operations designed to achieve the output (Tiryaki and Aydın, 2014).

Fig. 1.
The Study's Artificial Neural Network Structure

The output of the MLP can be expressed in mathematical form as per the following equation:

$$Y = g(\theta + \Sigma_{j=1}^{m} v_j + [\Sigma_{i=1}^{n} f(W_{ij} X_i + \beta_j)]) \quad (2)$$

In Equation (2), Y is the prediction value of dependent variable; X_i is the input value of ith independent variable; W_{ij} is the weight of connection between the ith input neuron and jth hidden neuron; β_j is the bias value of the jth hidden neuron; v_j is the weight of connection between the jth hidden neuron and output neuron; Θ is the bias value of output neuron; $g(.)$ and $f(.)$ are the activation functions of output and hidden neurons respectively (Tiryaki and Aydın, 2014).

3. Data Collection and Variable Selection

Air travel demand is influenced by a variety of exogenous and endogenous factors. Each factor is composed of elements that can stimulate or alternately reduce air travel growth. For passenger air traffic demand forecasting purposes, these factors are more conveniently categorised into two broad groups, those external to the airline industry and those within the airline industry itself (Ba-Fail *et al.,* 2000). Following an extensive review of the literature the following factors were selected as candidate variables in the ANN modelling: world GDP, world population growth, world jet fuel prices, world air fares (yield is the proxy variable

for travel cost), Australia's tourism attractiveness (as measured by Australia's recorded bed capacities at Australia's tourist accommodation establishments), Australia's unemployment rate, the Australian/United States dollar exchange rate, the number of outbound flights and Australia's unemployment size. Australia's outbound passengers are comprised of Australian nationals as well as foreign travelers. Australia's unemployment rate was included as a candidate variable as unemployment rates have been reported as being a determinant of air travel demand (Clark *et al.,* 2009; McKnight, 2010).

Three dummy variables were included in the modelling to control for the influence of the 2000 Sydney Olympic Games (Dummy 1), and the Commonwealth Games held in Melbourne from 15 to March 26, 2006 (Dummy 2). The first dummy variable (Dummy 1) controlled for the influence of the Olympic Games held in Sydney in 2000. The Olympic Games ran for 17 days (the opening ceremony was held on 15 September 2000). The Paralympics were also staged in Sydney over a 12-day period shortly after the conclusion of the Olympic Games (Madden, 2002). International airlines operated extra flights to Sydney, during the Sydney Olympic Games to accommodate extra passenger demand. Thus, the dummy variable reflecting the influence of the Olympic Games is equal to 1 for year 2000 and 0 otherwise.

The second dummy variable 9/11 (Dummy 2) controlled for the impact that the tragic events of 9/11 had on air travel demand. The dummy variable was equal to 1 for year 2001 and 0 otherwise.

The third dummy variable (Dummy 3) accounted for the impact of the Commonwealth Games held in Melbourne from 15 to March 26, 2006. The 2006 Melbourne Commonwealth Games was the largest sporting and community event held in Victoria's history (KPMG, 2006). International airlines carried participants and visitors to Melbourne so they could attend Commonwealth Games events. Therefore, the dummy variable reflecting the impact of the Commonwealth Games is equal to 1 for year 2006 and 0 otherwise.

Whilst the global airline industry was adversely impacted by the 2008 and 2009 global financial crisis, Australia's outbound air travel demand increased from 11.5 million in 2008 to 12.0 million in 2009 and to 13.3 million in 2010. Thus, a dummy variable was not included in the modelling.

The availability of a consistent data set allows the use of annual data for the period 1993 to 2016. The data used in the ANN model was sourced from a variety of sources. Data on world real GDP and world population growth were sourced from the International Monetary Fund (IMF). Australia's unemployment numbers and recorded bed capacities at Australia's tourist accommodation establishments are from the Australia Bureau of Statistics (ABS). The Australian/United States dollar data are from the Reserve Bank of Australia (RBA). The airfare data are from Boeing Commercial Airplanes. The data on Australia's outbound international airline passengers are from the Bureau of Infrastructure, Transport and Regional Economics (BITRE). World jet fuel prices (expressed in Australian dollars) were sourced from the US Energy Information

Administration (EIA). To convert collected data from current prices to real or constant prices, consumer price index at 2011 constant prices was used (Ba-Fail *et al.*, 2000).

Eleven variables were considered as input variables in the artificial neural network model: world real GDP, world passenger yields (proxy for air fares), world population size, Australia's unemployment numbers, Australia's tourist attractiveness, outbound flights from Australia, world jet fuel prices, the Australian/United States dollar exchange rate, and three dummy variables controlling for the influence of the 2000 Sydney Olympic Games, the impact of 9/11, and the Commonwealth Games held in Melbourne in March 2006.

2.3. Model Evaluation Goodness-of-Fit Measures

Goodness-of-fit (GOF) statistics are useful when comparing results across multiple studies. These statistics are also beneficial when researchers are examining competing models in a single study. In addition, GOF statistics provide feedback on the level of knowledge about the uncertainty involved in the phenomenon of interest (Washington *et al.*, 2011). Five measures were used in the present study: mean square error (MSE), the root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and correlation coefficient (R) (Kunt *et al.*, 2011; Tiryaki and Aydın, 2014).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(t_i - td_i)^2 \qquad (3)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - td_i)^2} \qquad (4)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left[\left|\frac{t_i - td_i}{t_i}\right|\right] \qquad (5)$$

$$MAPE = \frac{1}{N}\left(\sum_{i=1}^{N}\left[\left|\frac{t_i - td_i}{t_i}\right|\right]\right) \times 100 \qquad (6)$$

$$R = \frac{\sum_{i=1}^{N}(t_i - \bar{t})(td_i - \overline{td})}{\sqrt{\sum_{i=1}^{N}(t_i - \bar{t})^2 \cdot \sum_{i=1}^{N}(td_i - \overline{td})^2}} \qquad (7)$$

Where t_i is the measured values, td_i is the predicted values, N is the total number of data, and $(\overline{td})$ is the average of the predicted values (Tiryaki and Aydın, 2014, p. 104).

2.4. Artificial Neural Network Training and Testing

Training is the algorithmic process in the hidden neuron where parameter weights can be adjusted appropriately to forecast accurately. Among various training algorithms, the back-propagation is the most popular algorithm used (Claveria and Torra, 2014; Guo and Wong, 2013). The basic idea of this algorithm is to propagate a gradient of the transfer function back and compare actual output from output units with a target output, then re-adjust weights backward in the network. Weights are adjusted and repeated until the mean squared error (MSE) between network prediction and actual data is close to the target (Tiryaki and Aydın, 2014).

For the training process, artificial neural networks (ANNs) are separated into three data sets: training is used for model fitting and selection, testing is used for evaluating the model's forecasting ability and validation data sets to determine the end-point for the

training process to avoid model over fitting (Alekseev and Seixas, 2009; Shahin, 2013; Tiryaki and Aydın, 2014). In this study, the data was randomly divided into a 70:15:15 ratio (Garrido *et al.,* 2014; Kunt *et al.,* 2011; Tiryaki and Aydın, 2014). To adjust the synaptic weights and thresholds during the ANN training process, each presentation of the training samples is called the training epoch (Da Silva *et al.,* 2017).
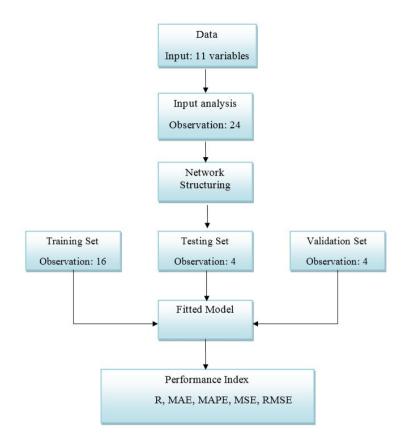
Insufficient training of the ANN will result in inaccurate results (Merilaita, 2010). Furthermore, over-fitting can be a concern with neural network model building (Grayson *et al.,* 2015; Merilaita, 2010; Myers *et al.,* 2016) as it can lead to predictions that are beyond the range of the training data (Richter and Weber, 2013). A cross validation process was carried out during the training phase to avoid model overfitting (Chew *et al.,* 2011; Efendigil *et al.,* 2009). Cross validation defines a method for testing the quality of the model during its learning phase, whilst at the same time avoiding over-training or over-fitting of the model. The principal aim of this process is to provide an indicator of the quality of the prediction model (Wiart, 2016, p. 141). The cross-validation data used in the study was comprised of randomly selected data which were separate to the model's training data (Chew *et al.,* 2011).

The objective of training is to minimize the global error such as root mean square error (RMSE), mean average error (MAE), mean square error (MSE), and mean absolute percent error (MAPE). ANNs usually commence with randomized weights for all their neurons. This means that they do not know everything and therefore require training to solve a particular problem for which they are intended. When a satisfactory level of performance is reached, training is

concluded and the network uses these weights to decide (Akgüngör and Doğan, 2009).

The training set was used to adapt the synaptic weights of the multilayer network, utilising the back propagation of estimation errors (Kalogirou, 2014). All inputs were inserted into the model and the networks trained. During the supervised learning process, an error function is defined. The synaptic weights values are iteratively updated until the provided output tends to be the desired, and the error function descends along the surface towards a local minimum. In this study, the training process stopped when it reached 1,000 epochs or 0.01 error tolerance (Efendigil *et al.,* 2009).

To conclude the training phase, a validation data set was used. The stopping criterion was the mean square error (MSE) of the estimated demand with respect to the samples belonging to the validation set. The validation set was not used in adapting the weight vectors of the neural estimator, and was therefore able to detect over-fitting in the training phase (Alekseev and Seixas, 2009).

For estimating the generalization capacity of the ANN forecasting models, a testing data set was also used (Da Silva *et al.,* 2017). Thus, after the training process was completed, a testing process was applied to ensure the model accuracy was sufficiently reliable. Once the values of the training set were determined, a data testing set was fed into the model and the output compared to the target value. The model was accepted if the difference was low enough (Garrido *et al.,* 2014). The testing set simulates the forecasting of the samples (Alekseev and Seixas, 2009). The study's artificial neural network (ANN) modelling process is presented in Figure 2.

Fig. 2.
The Study's Artificial Neural Network Modelling Process
Source: adapted from (Jiang et al., 2004, p. 7060).

2.5. Artificial Neural Network Transfer Function

The transfer function plays an important role in ANNs as it produces the output of the network. The transfer function or the activation in the hidden layer combines the inputs and weights values to deliver a signal to the output (Terzic *et al.,* 2012). This is usually a nonlinear function determining neuron output (Garrido *et al.,* 2014; Tiryaki and Aydın, 2014). The transfer or activation function typically falls into one of three categories:

- Linear (or ramp);
- Threshold;
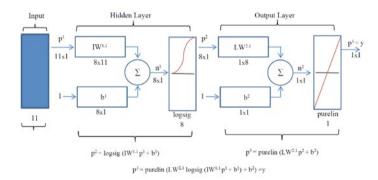- Sigmoid (Kar and Das, 2016; Terzic *et al.,* 2012).

The most frequently used transfer function is the sigmoid or logistic function as it possesses favourable mathematics properties such as montonicity, continuity, and

differentiability, which are all important when training a neural network with gradient descent (Priddy and Keller, 2005). An activation function is used as a boundary of output. These boundaries normally change from zero to one [0, 1] or from minus one to plus one [-1, +1] according to the type of activation function used in the ANN (Akgüngör and Doğan, 2009).

This study used the sigmoid function in the hidden layer and the linear transfer function in the output layer. The Levenberg–Marquardt back propagation algorithm was used as the training algorithm since its convergence is stable and fast (Ruiz-Aguilar *et al.,* 2014). The artificial neural network (ANN) tool box 8.0 within the framework of MATLAB R2012b (The MathWorks, Inc., USA.) software was used for modelling and simulation purposes.

3. Artificial Neural Network Modelling Results

3.1. Structure of the Final ANN Model for Predicting Australia's Outbound Air Travel Demand

The ANN model was developed to predict Australia's outbound air travel demand. The MLP model consisted of three layers having weight matrix W, bias vector b and output vector p^i where i >1. Figure 3 presents the optimum MLP model for predicting Australia's outbound air travel demand. The number of each layer is shown as a superscript to the variable of interest. Following (Kunt *et al.,* 2011), superscripts were used for identifying the source (second index) and destination (first index) for the various weights and other elements of the network.

$$p^2 = \text{logsig } (IW^{1,1} p^1 + b^1)$$

$$p^3 = \text{purelin } (LW^{2,1} p^2 + b^2)$$

$$p^3 = \text{purelin } (LW^{2,1} \text{logsig } (IW^{1,1} p^1 + b^1) + b^2) = y$$

Fig. 3.
The Structure of the Final Multi-Layer Perceptron Neural Network Model
Source: adapted from (Kunt et al., 2011, p.358)

The weight matrix connected to input vector p^1 was labelled as input weight matrix (IW^{1,1}). The elements of layer 1, such as its bias, net input and output have superscript 1 to indicate that they were associated with the first layer (Kunt *et al.,* 2011).

The matrices of layer weight (LW) and input weight (IW) were utilised in the MLP model. Data were randomly divided into three parts: training, testing, and validation (Alekseev and Seixas, 2009; Kunt *et al.,* 2011). The MLP model had 12 inputs, 9 neurons in the

hidden layers and 1 neuron in the output layer. The output layer of the MLP model consisted of one neuron representing Australia's outbound air travel demand. As noted earlier, 70 per cent of the data were used in the training phase. Validation and testing data sets each contained 15 per cent of the original data.

Constant input 1 was fed to the bias of each neuron. The outputs of each intermediate layer were the inputs to the subsequent layer. Hence, layer 2 can be analysed as one-layer having 8 inputs, 1 neuron and 1 X 8 weight matrix W^2. The layer can be treated as a single-layer network. The layers of a MLP play different roles in the prediction process (Kunt *et al.,* 2011). The back-propagation algorithm was applied to determine errors and modification for the weight of the hidden layer neurons (Akgüngör and Doğan, 2009). In this study, p^3 was the network output of interest and has been labelled as y (Rumelhart *et al.,* 1986).

The objective of this network is to reduce error e, which is the difference between t and p^i in which $i >1$ and t is the target vector. The perceptron learning rule calculates desired changes (target output) in the weights and biases of the perceptron, given input vector p^1 and the associated error e. Accordingly, the Least Mean Square Error (LMS) algorithm adjusts the weights and biases of the linear network to minimize this mean square error (Kunt *et al.,* 2011).

The error at output neuron j at iteration t can be calculated by the difference between the desired output (target output) and the corresponding real output, $e_j(t)=d_j(t)-y_j(t)$. So, Eq. (8) is the total error energy of all output neurons.

$$\varepsilon(t) = \tfrac{1}{2}\Sigma_{j\epsilon c}\, e_j^2\,(t) \tag{8}$$

Referring to Fig. 3, the output of the *k-th* neuron in the *l-th* layer can be calculated by Eq. (9) in which f_2 = log sig and f_3 = purelin:

$$y_k^l = f_k\left(\Sigma_{j=1}^{n^{l-1}} w_{jk}^l \cdot y_j^{l-1}\right) \tag{9}$$

where $1 \leq l \leq 3$, n^l refers to the number of neurons in layer l. For the input layer thus holds $l =1$, $y_j^1 = x_j$, for the output layer $l = 3$, $y_j^3 = y_j$.

The mean square error (MSE) of the output can be computed by:

$$E = \tfrac{1}{2}(d_j - y_j)^2 = \tfrac{1}{2}\left[d_j - f_3(\Sigma_{i=1}^9 w_{jk}^3 \cdot y_j^2)\right]^2 \tag{10}$$

The steepest descent of MSE can be used to update the weights by Eq. (11) (Yeung *et al.,* 2010):

$$w_{ij}^3(t+1) = w_{ij}^3(t) - \eta\,\frac{\partial E}{\partial w_{ij}^3} \tag{11}$$

The MSE performance index for the ANN is a quadratic function as shown in Eq. 10. Hence, the performance index will either have one global minimum, weak minimum or no minimum, depending upon the characteristics of input vectors (Kunt *et al.,* 2011). Specifically, the characteristics of input vectors determine whether a unique solution exists (Hagan *et al.,* 1996).

3.2. Final ANN Model Results

The final ANN PAX model in this study comprised 11 inputs, 8 neurons in the hidden layers and 1 neuron in the output layer. The final PAX ANN model is presented in the following equation

$$PAX = -0.326 + 0.12H_1 - 0.56H_2 - 0.61H_3 - 0.07H_4 + 0.15H_5 + 0.56H_6 + 0.44H_7 - 0.15H_8 \qquad (12)$$

Where: H_n = network hyperbolic tangent activation function:

$$H_n = TANH(Z_n) = \frac{(e^{Zn} - e^{-Zn})}{(e^{Zn} + e^{-Zn})} \qquad (13)$$

Where Z_n is calculated by multiplying the value of each input by the corresponding weight (w_i) (Eq.14) (Gonzalez, 2000).

$$Z_n = Bias_{Hn} + w_1 X_1 + w_2 X_2 + w_3 X_3 + w_4 X_4 + w_5 X_5 + w_6 X_6 + w_7 X_7 + w_8 X_8 + w_9 X_9 + w_{10} X_{10} + w_{11} X_{11} \qquad (14)$$

Where: X_1 = world real GDP; X_2 = world population; X_3 = World jet fuel prices; X_4 = World passenger yield; X_5 = Australia's tourist accommodation establishments; X_6 = Australia's outbound international flight numbers; X_7 = Australia's unemployment numbers; X_8 = Australian/United States dollar; X_9 = Dummy variable for Sydney Olympic Games; X_{10} = Dummy variable for 9/11; X_{11} = Dummy variable for Commonwealth Games.

The results of the PAX MLP model are presented in Table 1 in the form of a prediction table. Table 1 shows the prediction level of Australia's outbound air travel demand during training, testing, and validation phases.

Table 1
Prediction of the Outbound MLP Model
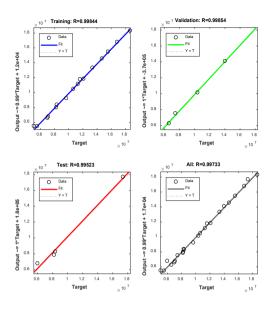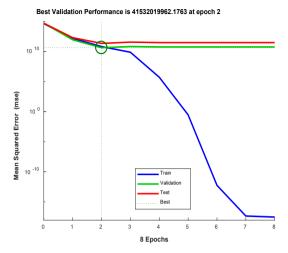
R	PAX Model
Training	0.99844
Validation	0.99854
Testing	0.99523
All	0.99733

Figure 4 shows the regression plots of the PAX model output with respect to training, validation and testing data. The value of the correlation coefficient (R) for each phase was also calculated (Kunt *et al.,* 2011). The R value was around 0.997 for the total response in the MLP model. The solid lines in Figure 4 shows the perfect linear fit between actual values and estimated values of Australia's outbound enplaned passengers. The correlation coefficient (R) between actual values and estimated values is another important indicator to check the validity of the model. Importantly, when the R value is close to 1, forecasting accuracy increases (Tiryaki and Aydın, 2014).
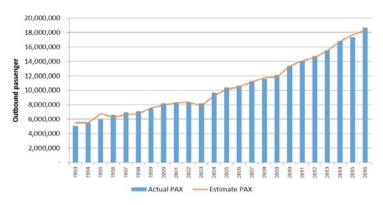
Fig. 4.

Regression Plots for Training, Testing and Validation Phases and the Total Response in the Outbound MLP Model

Training errors, validation errors and testing errors were plotted to determine validation errors in the training phase for the PAX model (Figure 5). The best validation performance in the PAX model occurred at epoch 2 with MSE at 4.1×10^{10} (Figure 5). The plot in Figure 5 shows the mean squared error commencing at a large value and decreasing to a smaller value, which indicates that network learning is improving. The plot in Figure 5 has three lines, because all input and target vectors were randomly divided into three sets (Garrido *et al.,* 2014; Kunt *et al.,* 2011). 70 per cent of the vectors were used for training the network. 15 per cent of the vectors were used for validating how well the network model was generalised. Training vectors continues for as long as it takes for training to reduce the network error on validation vectors. After the network has memorized the training set, training concludes. This technique automatically avoids the problem of over-fitting the model, which plagues many optimization and learning algorithms (Kunt *et al.,* 2011). As previously noted, the training process stopped when it reached 1,000 epochs or 0.01 error tolerance (Efendigil *et al.,* 2009). In order to estimate the generalization capacity of the neural PAX forecasting model, a testing set compromising the remaining 15 per cent of the vectors was used. This set was only presented to the neural estimator following conclusion of the training, and hence it did not participate in the training phase (Alekseev and Seixas, 2009).

Best Validation Performance is 41532019962.1763 at epoch 2

Fig. 5.
The Validation Error in the Outbound Air Travel Demand Model

Australia's actual and estimated outbound enplaned passengers during 1993 to 2016 are plotted and shown in Figure 6.

Fig. 6.
A Comparison of Australia's Actual and Estimated Outbound Air Travel Demand

This study used an artificial neural network (ANN) to predict Australia's outbound air travel demand. Table 2 presents mean absolute error (MAE), mean squared error, mean absolute percentage error (MAPE) and the root mean square error (RMSE) of the estimated model. These results suggest that the constructed ANN is promising for modelling Australia's outbound air travel demand.

Table 2

The Final Results of the Objective Function in Australia's Outbound Air Travel Demand ANN Model

Error	ANN
MAE	2.3×10^5
MSE	8.6×10^{11}
MAPE	2.81 %
RMSE	2.9×10^5

3.3. Discussion of Contributing Factors that Influence Australia's Outbound Air Travel Demand

This study used a contribution table (Gately, 1996) to analyse the major contributing factors that influence Australia's outbound airline passenger demand. The contribution of factor (C_i) in the input layer is the sum of absolute values of the weight of connection between the input neuron and the hidden neuron.

$$C_i = \sum_{j=1}^{k} \left| W_{ij} \right| \qquad (15)$$

where:
C_i is the contribution value of factor i;
W_{ij} is the weight of connection between the ith input neuron and jth hidden neuron.

The contributing factor scale developed by (Gately, 1996) was used to evaluate the influences of input variables. Based on this scale, any input variable with a contribution value less than 2 is considered a weak contributing factor while any input variable with a contribution value greater than 5 is considered a high contributing factor (Chen *et al.,* 2012).

Table 3 shows the contribution value of input variables in the PAX model and shows that all the input variables in the PAX model have a contribution value higher than 2 which means that no input variables are considered a weak contributing factor. Also, the four most important input variables for forecasting Australia's outbound air travel demand are: X_2= World Population, X_5 = Australia's tourist accommodation establishments, X_8 Australian/United States dollar exchange rate, and X_{10} Dummy variable for 9/11.

Table 3

The Contributions of the Study's Input Variables

Input Variables	PAX Model
X_1 = world real GDP	2.77
X_2 = world population	4.16
X_3 = World jet fuel prices	2.76
X_4 = World passenger yield	3.42
X_5 = Australia's tourist accommodation establishments	4.95
X_6 = Australia's outbound international flight numbers	3.53
X_7 = Australia's unemployment numbers	2.55
X_8 = Australian/United States dollar	4.57
X_9 = Dummy variable for Sydney Olympic Games	2.01
X_{10} = Dummy variable for 9/11	4.21
X_{11} = Dummy variable for Commonwealth Games	3.42

4. Conclusions

In the global airline industry, forecasting is regarded as one of the most critical areas of airline management. The aim of this study was for the first time to develop an artificial neural network (ANN) models for predicting Australia's outbound international airline passenger demand. Based on the use of annual data from 1993 to 2016, the ANN model was developed and empirically tested using the input parameters of world GDP, world population, world jet fuel prices, world air fares (proxy for air travel cost), Australia's tourism attractiveness, outbound flights, Australia's unemployment levels, the Australian and United States foreign exchange rate and three dummy variables. The dummy variables were incorporated to control for the influence of the Sydney Olympic Games in 2000, the tragic events of 9/11 and the Commonwealth Games, which were held in Melbourne in 2006. The artificial neural network (ANN) used multi-layer perceptron (MLP) architecture that compromised a multi-layer feed-forward network and the sigmoid and linear functions were used as activation functions with the feed forward-back propagation algorithm. The ANN was applied during training, testing and validation and had 8 inputs, 1 neurons in the hidden layers and 1 neuron in the output layer. The data was randomly divided into three data sets. 70% of the data was used in the training phase with the remaining data divided into validation (15%) and testing (15%). To identify the best-fit model, four goodness-of-fit measures: mean absolute error (MAE), mean square error (MSE), root mean square errors (RMSE), and mean absolute percentage errors (MAPE) were used in the study. The highest R-value obtained from the ANN model is 0.99733, demonstrating that the ANN provided a high predictive capability.

Following the recommendations of (Gately, 1996), this study used a contribution table to analyse the factors that influence Australia's outbound air travel demand. The input variables with the highest contributing factor were Australia's tourism attractiveness (4.95), the Australian/United States dollar exchange rate (4.57), world population (4.16) and the impact of 9/11 (Dummy variable X_{10}) (4.21). Australia has long been a major tourism destination and throughout the study's review period there has been a significant growth in tourists from China, Japan, New Zealand and the USA. Population has a direct effect on the size of an air travel market and may cause a bias in the estimates if omitted. For instance, a large increase in air traffic may reflect a sudden increase in population rather than other effects (International Air Transport Association, 2008). The months following the terror attacks of 9/11, had a significant impact of the demand for air travel.

References

Abed, S.Y.; Ba-Fail, A.O.; Jasimuddin, S.M. 2001. An econometric analysis of international air travel demand in Saudi Arabia, *Journal of Air Management* 7(3): 143-148.

Aderamo, A.J. 2010. Demand for air transport in Nigeria, *Journal of Economics* 1(1): 23-31.

Akgüngör, A.P.; Doğan, E. 2009. An artificial intelligent approach to traffic accident estimation: model development and application, *Transport* 24(2): 135–142.

Alekseev, K.P.G.; Seixas, J.M. 2002. Forecasting the air transport demand for passengers with neural modelling, in *Proceedings of the VII Brazilian Symposium on Neural Networks (SBRN/02)*, 11–14 November 2002, Recife, Brazil.

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(5): 212–216.

Ba-Fail, A.O. 2004. Applying data mining techniques to forecast number of airline passengers in Saudi Arabia (domestic and international travels), *Journal of Air Transportation* 9(1): 100-115.

Ba-Fail, A.O.; Abed, S.Y.; Jasimuddin, S.M. 2000. The determinants of domestic air travel demand in the Kingdom of Saudi Arabia, *Journal of Air Transportation World Wide* 5(2): 72–86.

Bureau of Infrastructure, Transport and Regional Economics. 2017. International airline activity, Statistical Report. Available from internet: < https://bitre.gov.au/publications/ongoing/files/International_airline_activity_CY2016.pdf>.

Bhadra, D. 2003. Demand for air travel in the United States: bottom-up econometric estimation and implications for forecasts by origin and destination pairs, *Journal of Air Transportation* 8(2):19–56.

Chen, S.C.; Kuo, S.Y.; Chang, K.W.; Wang, Y.T. 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.

Chew, E.P.; Lee, L.H.; Tan, L.C. 2011. *Advances in maritime logistics and supply chain systems*. World Scientific Publishing, Singapore. 332 p.

Clark, D.E.; McGibany, J.M.; Myers, A. 2009. The effects of 9/11 on the airline travel industry. In book (ed. Morgan M.J.) *The Impact Of 9/11 On Business and Economics: The Business of Terror - The Day That Changed Everything*? Palgrave Macmillan, United Kingdom: 76-86.

Claveria, O.; Torra, S. 2014. Forecasting tourism demand to Catalonia: neural networks vs. time series models, *Economic Modelling* 36: 220-228.

Da Silva, I.N.; Spatti, D.H.; Flauzino, R.A.; Liboni, L.H.B.; dos Reis Alves, S.F. 2017. *Artificial neural networks: a practical course.* Springer International Publishing, Switzerland. 307 p.

Dulikravich, G.S.; Colaço, M.J. 2015. Hybrid optimization algorithms and hybrid response surfaces. In book (*ed. Greiner, D. et al.*) *Advances in Evolutionary and Deterministic Methods for Design, Optimization and Control in Engineering Sciences*, Springer International Publishing, Switzerland: 19-48.

Efendigil, T.; Önüt, S.; Kahraman, C. 2009. A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: a comparative analysis, *Expert Systems with Applications* 36(3): 6697-6707.

Garrido, C.; De Oña, R.; De Oña, J. 2014. Neural networks for analyzing service quality in public transportation, *Expert Systems with Applications* 41(15): 6830–6838.

Gately, E. 1996. *Neural networks for financial forecasting*. John Wiley & Sons, Inc, United States. 169 p.

Gonzalez, S. 2000. Neural Networks for Macroeconomic Forecasting: A Complementary Approach to Linear Regression Models, Finance Canada Working Paper 2000-07. Available from internet: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.323.8403&rep=rep1&type=pdf>.

Grayson, J.; Gardner, S.; Stephens, M. 2015. *Building better models with JMP Pro*. SAS Institute, Inc, United States. 352 p.

Guo, Z.X.; Wong, W.K. 2013. Fundamentals of artificial intelligence techniques for apparel management applications. In book: (*ed. Wong, WK. et al.*) *Optimizing Decision Making in the Apparel Supply Chain Using Artificial Intelligence (AI): From Production to Retail*. Woodhead Publishing, United Kingdom: 13-40.

Hagan, M.T.; Demuth, H.B.; Beale, M.H. 1996. *Neural network design*, PWS Publishing, Boston. 800 p.

Jiang, D.; Zhang, Y.; Hu, X.; Zeng, Y.; Tan, J.; Shao, D. 2004. Progress in developing an ANN model for air pollution index forecast, *Atmospheric Environment* 38(40): 7055–7064.

International Air Transport Association. 2008. Air travel demand. IATA Economics Briefing No. 9. Available from Internet: <http://www.iata.org/ whatwedo/Documents/economics/air_travel_demand.pdf>.

Kale, M.U.; Deshmukh, M.M.; Wadatkar, S.B.; Talokar, A.S. 2016. Constraints in rainfall-runoff using artificial neural network. In book (ed. Panigrahi, B., Goyal, M.R.) *Modeling Methods and Practices in Soil and Water Engineering*, Apple Academic Press, Inc, Canada: 31-40.

Kalogirou, S.A. 2014. *Solar energy engineering: processes and systems*. Second Edition, Academic Press, United Kingdom. 840 p.

Kar, P.; Das, A. 2016. Artificial neural networks and learning techniques. In book (ed. Samui, P.) *Handbook of Research on Advanced Computational Techniques for Simulation-Based Engineering*, IGI Global, United States: 227-251.

Kopsch, F. 2012. A demand model for domestic air travel in Sweden, *Journal of Air Transport Management* 20: 46–48.

KPMG. 2006. Economic Impact Study of the Melbourne 2006 Commonwealth Games Post-Event Analysis, Office of Commonwealth Games Coordination. Available from internet: <https://opus.lib.uts.edu.au/bitstream/10453/19802/1/econ_impact_report.pdf>.

Kunt, M.M.; Aghayan, I.; Noii, N. 2011. Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods, *Transport* 26(4): 353–366.

Luxton, D.D. 2016. An introduction to artificial intelligence in behavioral and mental healthcare. In book (ed. Loxton, D.D.) *Artificial Intelligence in Behavioral and Mental Health Care*, Academic Press, United Kingdom: 1-26.

Madden, J.R. 2002. The economic consequences of the Sydney Olympics: the CREA/Arthur Andersen study, *Current Issues in Tourism* 5(1): 7–21.

McKnight, P. 2010. Airline economics. In book (ed. Wald A., Fay C., Gleich R.) *Introduction to Aviation Management*, LIT Verlag, Germany: 25-53.

Merilaita, S. 2010. Applying artificial neural networks to the study of prey colouration. In book (ed. Tosh, C.R., Ruxton, G.D.) *Modelling Perception with Artificial Neural Networks*, Cambridge University Press, United Kingdom: 215-235.

Merkus, H.G.; Meesters, G.M.H. 2014. Introduction. In book (ed. Merkus, H.G., Meesters, G.M.H.) *Particulate Products: Tailoring Properties for Optimal Performance*, Springer International Publishing, Switzerland: 1-19.

Myers, R.H., Montgomery, D.C.; Anderson-Cook, C.M. 2016. *Response surface methodology: process and product optimization using design experiments*. Forth Edition, John Wiley & Sons, United States. 856 p.

Priddy, K.L.; Keller, P.E. 2005. *Artificial neural networks: an introduction*. SPIE Press, United States. 180 p.

Rao, R.K. 2011. *Advanced modeling and optimization of manufacturing processes: international research and development*. Springer Verlag London Ltd, United Kingdom. 380 p.

Remennikov, A.M.; Mendis, P.A. 2016. Prediction of airblast loads in complex environments using artificial neural networks. In book (ed. Syngellakis, S.) *Design Against Blast: Load Definition & Structural Response*. WIT Press, United Kingdom: 53-62.

Richter, M.M.; Weber, R.O. 2013. *Case-based reasoning: a textbook*. Springer-Verlag, Germany. 546 p.

Ruiz-Aguilar, J.J.; Turias, I.J.; Jiménez-Come, M.J. 2014. Hybrid approaches based on SARIMA and artificial neural networks for inspection time series forecasting', *Transportation Research Part E: Logistics and Transportation Review* 67: 1–13.

Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. 1986. *Parallel distributed processing: explorations in the microstructure of cognition: foundations*. The MIT Press, United States. 567 p.

Santos, A. A. P.; Junkes, L.N.; Pires Jr, F.C.M. 2014. Forecasting period charter rates of VLCC tankers through neural networks: a comparison of alternative approaches, *Maritime Economics & Logistics* 16(1): 72–91.

Sen, J.; Sas, A.K. 2014. Artificial neural network model for forecasting the stock price of Indian IT company. In book (*ed. Babu, B.V. et al.*) *Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012*, December 28-30. Springer India, India: 1151-1160.

Shahin, M.A. 2013. Artificial intelligence in geotechnical engineering: applications, modeling aspects, and future directions. In Book (*ed. Yang, X.S. et al.*) *Metaheuristics in Water, Geotechnical and Transport Engineering*. Elsevier, United Kingdom: 169-204.

Shirgure, P.S.; Rajput, G.S. 2014. Evaporation estimations with neural networks. In book (ed. Goyal, M.R., Harmsen, E.W.) *Evapotranspiration: Principles and Applications for Water Management*. Apple Academic Press, Canada, 131-164.

Sineglazov, V.; Chumachenko, E.; Gorbatyuk, V. 2013. An algorithm for solving the problem of forecasting, *Aviation* 17(1): 9–13.

Sivanandam, S.N.; Sumathi, S.; Deepa, S.N. 2006. *Introduction to neural networks using Matlab 6.0*. Tata McGraw-Hill, India. 656 p.

Sivrikaya, O.; Tunç, E. 2013. Demand forecasting for domestic air transportation in Turkey, *The Open Transportation Journal* 7: 20–26.

Skias, S.T. 2006. Background of the verification and validation of neural networks. In book (ed. Taylor, B.J.) *Methods and Procedures for the Verification and Validation of Artificial Neural Networks*. Springer Science+Business Media, United States: 1-12.

Srisaeng, P.; Baxter, G.; Wild, G. 2015a. Using an artificial neural network approach to forecast Australia's domestic passenger air travel demand, *World Review of Intermodal Transportation Research* 5(3): 281-313.

Srisaeng, P.; Baxter, G.S.; Wild, G. 2015b. Forecasting demand for low cost carriers in Australia using an artificial neural network approach, *Aviation* 19(2): 90-103.

Terzic, E.; Terzic, J.; Nagarajah, R.; Alamgir, M. 2012. *A neural network approach to fluid quantity measurement in dynamic environments*. Springer-Verlag, United Kingdom. 140 p.

Tiryaki, S.; Aydın, A. 2014. An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model, *Construction and Building Materials* 62: 102–108.

Washington, S.P.; Karlaftis, M.G.; Mannering, F. 2011. *Statistical and econometric methods for transportation data analysis*. Second Edition, Chapman & Hall/CRC Press, United States. 544 p.

Wiart, J. 2016. *Radio-frequency human exposure assessment: from deterministic to stochastic methods*. John Wiley & Sons, United States. 196 p.

Yang, J.L.; Ma, J.; Howard, S.K. 2016. A structure optimization algorithm of neural networks for pattern learning from educational data. In book (ed. Shanmuganathan, S., Sandhya, S) *Artificial Neural Network Modelling*. Springer International Publishing, Switzerland: 67-82.

Yeung, D.S. et. al. 2010. *Sensitivity analysis for neural networks*. Springer, Germany. 86 p.