AN ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM FOR FORECASTING AUSTRALIA’S DOMESTIC LOW COST CARRIER PASSENGER DEMAND

Panarat SRISAENG1, Glenn S. BAXTER2, Graham WILD3

School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University, Melbourne, Australia 3001
E-mail: 1s3125221@student.rmit.edu.au (corresponding author); 2glenn.baxter@rmit.edu.au; 3graham.wild@rmit.edu.au

Received 7 April 2014; accepted 12 February 2015

Panarat SRISAENG, PhD
Education: Bachelor of Economics, Chulalongkorn University, Bangkok, Thailand, 1993. Master of Business Economics, Kasetsart University, Bangkok, Thailand, 1998. Affiliations and functions: PhD Aviation, RMIT University, School of Aerospace, Mechanical and Manufacturing Engineering and is a lecturer in airline business management at Suan Dusit University, School of Tourism and Hospitality Management, Thailand. Research interests: low cost airlines management, demand model for air transportation, demand forecasting for air transportation.

Glenn S. BAXTER, PhD
Education: Bachelor of Aviation Studies, University of Western Sydney, Australia, 2000. Master of Aviation Studies, University of Western Sydney, Australia, 2002. PhD, School of Aviation, Griffith University, Brisbane, Australia, 2011. Affiliations and functions: Lecturer in Aviation Management and Deputy Aviation Programs Director, RMIT University, School of Aerospace, Mechanical and Manufacturing Engineering. Research interests: air cargo handling and operations, airport operations and sustainability, supply chain management.

Graham WILD, PhD
Abstract. This study has proposed and empirically tested two Adaptive Neuro-Fuzzy Inference System (ANFIS) models for the first time for predicting Australia’s domestic low cost carriers’ demand, as measured by enplaned passengers (PAX Model) and revenue passenger kilometres performed (RPKs Model). In the ANFIS, both the learning capabilities of an artificial neural network (ANN) and the reasoning capabilities of fuzzy logic are combined to provide enhanced prediction capabilities, as compared to using a single methodology. Sugeno fuzzy rules were used in the ANFIS structure and the Gaussian membership function and linear membership functions were also developed. The hybrid learning algorithm and the subtractive clustering partition method were used to generate the optimum ANFIS models. Data was normalized in order to increase the model’s training performance. The results found that the mean absolute percentage error (MAPE) for the overall data set of the PAX and RPKs models was 1.52% and 1.17%, respectively. The highest $R^2$-value for the PAX model was 0.9949 and 0.9953 for the RPKs model, demonstrating that the models have high predictive capabilities.

**Keywords:** adaptive neuro-fuzzy inference system (ANFIS); air transport; Australia; forecasting methods; low cost carriers.

1. Introduction

Forecasting is the process of making projections about future performance based on the existing historic data. An accurate forecast assists firms in decision-making and planning for the future. Forecasts empower people to modify existing variables at the current time to predict the future in order to achieve a favorable scenario (Hadavandi et al. 2011).

Forecasting passenger transport demand is of critical importance for airlines as well as for investors since investment efficiency is greatly influenced by the accuracy and adequacy of the estimation performed (Blinova 2007). Air traffic forecasts are therefore one of the key inputs into an airline’s fleet planning and route network development, and are also used in the preparation of the airline’s annual operating plan (Ba-Fail et al. 2000; Doganis 2009). Furthermore, analysing and forecasting air travel demand may also assist an airline in reducing its risk through an objective evaluation of the demand side of the airline business (Abed et al. 2001; Ba-Fail et al. 2000).

Classical modelling such as multiple linear regression (MLR) has been used extensively in forecasting air traffic demand for several decades (see, for example, Abed et al. 2001; Aderamo 2010; Ba-Fail et al. 2000; International Civil… 2006; Kopsch 2012; Sivrikaya, Tunç 2013). However, traditional regression techniques are not able to capture the non-linear structure of a specific process as effectively as the artificial intelligence-based models. Consequently, artificial intelligence-based modelling techniques have become more popular in diverse disciplines over the past decade (Kar et al. 2014) because of their robustness, high predictive capabilities and flexible behaviours to handle the multi-objective criteria in a straightforward manner (Yetilmeszoy et al. 2011).

Jang (1993) and Jang et al. (1997) introduced the adaptive network-based fuzzy inference system (ANFIS), which is a system using a hybrid learning rule to optimize the fuzzy system parameters of a first order Sugeno system (Giovanis 2012). This approach has been applied to a growing range of disciplines, including transport mode choice (Andrade et al. 2007), economics (Fang 2012; Giovanis 2012), electricity demand forecasting (Zahedi et al. 2013), financial markets forecasting (Bagheri et al. 2014; Kablan 2009), gold price forecasting (Makridou et al. 2013), oil consumption forecasting (Senvar et al. 2013), stock market forecasting (Atsalakis, Valavanis 2009; Chen et al. 2013; Cheng et al. 2013; Svalina et al. 2013; Wei 2013), tourism demand forecasting (Atsalakis et al. 2014; Chen et al. 2010; Hadavandi et al. 2011), and ordering policy in supply chains (Latif et al. 2014).

Following the deregulation of Australia’s domestic airline market in 1990, which permitted other airlines to compete with the established carriers (Forsyth 2003; Nolan 1996), a number of low cost carriers (LCCs) have entered the market. The LCCs now hold around a 35 per cent market share, with the two major incumbent LCCs being “Jetstar Airways” and “Tiger Airways” (Srisaeng et al. 2014). Despite the reported advantages of the ANFIS in the literature together with acknowledged critical importance of forecasting for airline and airport management, to the best of the author’s knowledge there has been no previously reported study that has developed and empirically examined ANFIS models for forecasting Australia’s domestic quarterly LCCs air travel demand and revenue passenger kilometres performed (RPKs). Thus, the key objective of this study is to address this apparent research gap in the literature. Furthermore, the study is intended to provide both a theoretical perspective on the development of an ANFIS for forecasting airline passenger demand, and also to provide a practical application of the ANFIS to forecast Australia’s quarterly domestic LCC passenger demand and revenue passenger kilometres performed (RPKs).

---

1 Enplaned passengers and revenue passenger kilometres performed (RPKs) are the recognised measures of airline passenger traffic (Belobaba 2009; Holloway 2008).
The remainder of the paper is structured as follows: Section 2 presents the literature review and this is followed by the ANFIS architecture in Section 3. The proposed ANFIS method and a real-world case study focusing on forecasting Australia’s LCCs passenger demand is presented in Section 4. Section 5 contains some concluding remarks.

2. Adaptive neuro-fuzzy inference system (ANFIS)

The Adaptive Neuro-Fuzzy Inference System (ANFIS), first introduced by Jang (1993), is a hybrid method comprising both fuzzy inference systems and the artificial neural network (ANN) (Fang 2012; Liu et al. 2008). This system therefore combines the benefits of both approaches; wherein the former brings prior knowledge into a set of constraints to obtain the optimal solution, the latter is good at capturing various patterns (Jang et al. 1997; Xiao et al. 2014; Yetilmezsoy et al. 2011). An ANFIS’s principal objective is the determination of the optimum values of equivalent fuzzy inference system parameters. This is achieved through the application of a learning algorithm using input–output data sets. The optimisation of parameters during the training session is undertaken in such a way that the error between the target and actual output is minimized (Goyal et al. 2014). The parameters to be optimized in ANFIS are the premise parameters, which describe the shape of the membership functions, and the consequent parameters, which describe the overall output of the ANFIS system. The optimum parameters obtained are then used in the testing session to calculate the prediction (Mayilvaganan, Naidu 2011).

The ANFIS is considered a more powerful approach than the simple fuzzy logic algorithm and artificial neural networks, as this technique provides a method whereby fuzzy modelling learns about the data set; in order to compute the membership function parameters which best allow the associated fuzzy inference system to track the given input/output data (Al-Ghandoor et al. 2012: 130). A further advantage of the ANFIS is the fact that it can be trained without the requirement for the expert knowledge normally required for the standard fuzzy logic design, and both numerical and linguistic knowledge can be combined into a fuzzy rule base by utilising fuzzy methods (Giovanis 2012). Other important advantages of the ANFIS include its nonlinear ability, its capacity for rapid learning, and its adaptation capability. Furthermore, the strength of the ANFIS is that it uses the artificial neural network’s ability to classify data and identify patterns. Moreover, the ANFIS develops a fuzzy expert system that is more transparent to the user and which is also less likely to produce memorization errors than an ANN (Giovanis 2012).

Artificial neural network-based methods have been used successfully for modelling across a broad range of disciplines (Yetilmezsoy et al. 2011). However, poor interpretation has been reported as a major drawback of their utilization (Wieland et al. 2002). A major shortcoming of artificial neural networks (ANNs) is that they are unable to reveal causal relationships between major system components. Consequently, they are unable to improve the user’s explicit knowledge (Yetilmezsoy et al. 2011). Therefore, to overcome the problematic conditions of ANNs and fuzzy systems, a new system combining both ANNs and the fuzzy system, called the adaptive-network-based fuzzy inference system (ANFIS) was proposed by Jang (1993).

3. ANFIS architecture

The ANFIS is an adaptive network compromising nodes and directional links with associated learning rules. It is called adaptive because some, or all, of the nodes have parameters which influence the output of the node. The ANFIS identifies and learns the relationships between inputs (Kablan 2009). The operation of the ANFIS resembles the feed forward back propagated (FFBP) artificial neural network. Consequent parameters are calculated forward, while premise parameters are calculated backward. The ANFIS comprises two parts, the antecedent and the conclusion, and these are connected to each other by fuzzy rules based on the network form (Yetilmezsoy et al. 2011). There are two types of fuzzy inference systems: Mamdani-type and Sugeno-type. These two types of fuzzy inference systems do, however, vary somewhat in the ways that the outputs are determined. The principal difference between the two FIS types is that, in the Sugeno-system, the output membership functions are either constant or linear (Arkhipov et al. 2008: 496).

This study used the Sugeno-type FIS system. There are two learning methods in the system’s neural section: a hybrid learning method and a back propagation (BP) learning method (Yetilmezsoy et al. 2011). The output variables are obtained by applying fuzzy rules to fuzzy sets of input variables (Cakmakci et al. 2010; Jang 1993; Takagi, Sugeno 1985). A typical ANFIS employs a Takagi-Sugeno model-based fuzzy inference approach in order to form the related hybrid system (Köse, Arslan 2013). To present the ANFIS architecture, two fuzzy if-then rules based on a first order Sugeno model are considered (Bagheri et al. 2014; Übeyli et al. 2010).

Rule 1: if x is A1 and y is B1, then:
\[ f_1 = p_1 x + q_1 y + r_1. \]

Rule 2: if x is A2 and y is B2, then:
\[ f_2 = p_2 x + q_2 y + r_2, \]

where x and y are the inputs, A_i and B_i are the fuzzy sets, f_i are the outputs within a fuzzy rule, and p_i, q_i, and r_i
are the design parameters which are determined during the training process (Übeyli et al. 2010). Figure 1 shows the fuzzy reasoning mechanism (Takagi, Sugeno 1983).

The ANFIS architecture used to implement these two rules is depicted in Figure 2. The ANFIS architecture comprises 5 layers, that is, a fuzzy layer, a product layer, a normalised layer, a defuzzy layer and a total output layer (Jang 1993; Yetilmezsoy et al. 2011). As shown in Figure 2, each node in the ANFIS is characterized by a node function with fixed or adjustable parameters; a circle indicates a fixed node, whereas a square indicates an adaptive node (Ch, Mathur 2010). Model parameter values are determined through the learning or training phase of an artificial neural network. The model performance is evaluated by the satisfactorily fitted training and test data. Furthermore, the model performance evaluates error values, for example, the root mean square error (RMSE), which is minimized in turn through backpropagation as well as the hybrid learning algorithms allowed by the ANFIS.

Each layer of the ANFIS has its own task, so the following section describes the relationship between the output and input layer in the ANFIS.

Layer 1 is the fuzzification layer that passes crisp external signals directly to the following layer (Xiao et al. 2014). In the fuzzy layer, x and y are the input of nodes $A_1$, $A_2$, $B_1$, and $B_2$, respectively. $A_1$, $A_2$, $B_1$, and $B_2$ are the linguistic labels used in fuzzy theory for dividing membership functions (Yetilmezsoy et al. 2011). Every node $i$ in layer 1 is an adaptive node which has a specific function (Übeyli et al. 2010; Yetilmezsoy et al. 2011). The nodes in layer 1 implement the fuzzy membership functions and map the input variables to the corresponding fuzzy membership values (Yetilmezsoy et al. 2011). The parameters in this layer are called premise parameters (Yilmaz, Kaynar 2011). The output of layer 1 indicates the degree/grade of the fuzzy membership function of the given inputs which are determined by the fuzzy membership function (Xiao et al. 2014). The output of layer 1 is:

$$O_i^1 = \mu_{A_i}(x), i = 1, 2 \quad \text{or} \quad O_i^2 = \mu_{B_i}(y), i = 3, 4 , \quad (1)$$

where $x$ and $y$ are the input to the $i$th node and $A_i$ and $B_{i,2}$ are the linguistic labels associated with this node (Xiao et al. 2014).

Thus, $O_i^1$ is the membership grade of a fuzzy set $A_i$ (=$A_1$, $A_2$, $B_1$, or $B_2$) and it specifies the degree to which the given input $x/y$ satisfies the quantifier $A_i$, where $\mu_{A_i}(x)$ and $\mu_{B_{i,2}}(y)$ can adopt any fuzzy membership function, for example, if the bell shaped membership function is employed, as given by (Übeyli et al. 2010).

$$\mu_{A_i}(x) = \frac{1}{1 + \left[\frac{x-a_i}{b_i}\right]^n} , \quad (2)$$

where {a$_i$, b$_i$, c$_i$} are the parameters of the function (Xiao et al. 2014). During the learning stage the back propagation algorithm adopts their values. As the values of these parameters change, the bell-shape function varies accordingly, hence exhibiting various forms of membership functions on linguistic label $A_i$ (Xiao et al. 2014; Yetilmezsoy et al. 2011).

Layer 2 is a rule layer, each node is a representation of a rule and the inputs are the degrees of membership functions which are multiplied through a T-norm operator so as to determine the level of fulfilment of $W_i$, the rule (Ch, Mathur 2010). The nodes are fixed nodes and are labelled “∏”, which indicates that they perform as a single multiplier (Übeyli et al. 2010). Each node represents the firing strength of the reasoning rule (Patil et al. 2011; Yilmaz, Kaynar 2011). The outputs of this layer can be represented as:

$$O_i^2 = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), i = 1, 2 . \quad (3)$$
Layer 3 is the normalization layer, whose nodes are labelled “N”, indicating that they play a normalization role to the firing strengths from the previous layer (Übeyli et al. 2010; Yetilmezsoy et al. 2011). This layer normalizes each rule’s output with respect to the rest of the rule set, and the normalization scales the rule’s output to a value between zero and one by dividing its output by the number of inputs (Schott, Kalita 2011).

\[ O^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2, \tag{4}\]

where \( w_i \) is the firing strength of the \( i \)th rule which is computed in layer 2. Node \( i \) computes the ratio of the \( i \)th rule’s firing strength to the sum of all rules’ firing strengths (Xiao et al. 2014).

Layer 4 is the defuzzification layer in which the nodes are adaptive nodes (Übeyli et al. 2010; Yetilmezsoy et al. 2011). Every node in layer 4 computes a linear function, where function coefficients are adapted by using the error function of the multilayer feedforward neural network (Xiao et al. 2014). The parameters in this layer are called consequent parameters (Yilmaz, Kaynar 2011).

\[ O^4 = \bar{w}_i f_i = \bar{w}_i \left( p_1 x + q_1 y + n_1 \right), \quad i = 1, 2, \tag{5}\]

\( \left( p_1, q_1, n_1 \right) \) are the parameter set.

The fifth ANFIS layer, whose node is labelled “∑”, is the output layer, in which a single node calculates the overall output as a summation of all incoming signals (Ch, Mathur 2010; Giovannis 2012; Xiao et al. 2014). Hence, the overall output of the model can be written as (Fang 2012; Yetilmezsoy et al. 2011):

\[ O^5 = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}, \tag{6}\]

where \( \bar{w}_i f_i \) denotes the consequent part of rule \( i \). The overall output of the neuro-fuzzy system is the summation of the rule consequences (Xiao et al. 2014).

As previously noted, in the ANFIS structure, the premise and consequent parameters are important factors for the learning algorithm in which each parameter is used to calculate the output data of the training data (Efendigil et al. 2009). The premise part of a rule defines a subspace, whereas the consequent part specifies the output within this fuzzy subspace (Jang 1993).

The ANFIS allows for the use of two learning algorithms, back propagation and hybrid methods, which seek to minimize some measure of error, for instance, the root mean sum of squared differences (RMSE) between the observed and predicted data (Yetilmezsoy et al. 2011). The hybrid learning rule combines the gradient method and the least squares estimate to identify optimal parameters (Jang 1993). For the hybrid learning algorithm, it can be noted that when the values of the premise parameters are fixed, the overall output can then be expressed as a linear combination of the consequent parameters (Yetilmezsoy et al. 2011).

Although ANN and fuzzy-logic models are both basic areas of the artificial intelligence concept (Yetilmezsoy et al. 2011), the ANFIS combines and captures the advantages of both of these methods (Ch, Mathur 2010; Liu et al. 2008; Tiwari et al. 2012). Since the ANFIS is an adaptive network which permits the use of ANN topology as well as fuzzy logic, it includes the characteristics of both and also eliminates some disadvantages of these methods when used individually (Yetilmezsoy et al. 2011). Accordingly, the ANFIS is capable of handling complex and nonlinear problems (Giovannis 2012). Even if the targets are not given, the ANFIS may reach the optimum result very quickly (Yetilmezsoy et al. 2011). Furthermore, there is no vagueness in the ANFIS as opposed to ANNs (Tiwari et al. 2012). This implies that the ANFIS may reach the target more rapidly than ANNs (Kumar et al. 2011). Therefore, when a more sophisticated system with high-dimensional data is implemented, the use of the ANFIS instead of the ANN is considered to be more appropriate to overcome the complexity of the problem faster (Noori et al. 2009).

In the ANFIS structure, the implication of the errors is different from that of ANNs. In order to ascertain an optimal result, the epoch size is not limited (Noori et al. 2009). In training high-dimensional data, the ANFIS can provide results with the minimum total error as compared to ANN and fuzzy logic methods (Chi et al. 2005; Yetilmezsoy et al. 2011).

In the ANFIS system, each input parameter may be clustered into several class values in layer 1 to build up fuzzy rules. Each fuzzy rule would be constructed using two or more membership functions in layer 2. Several methods have been proposed to classify the input data and for the rule-making, among which the most common being the grid partition (Jang et al. 1997) and the subtractive fuzzy clustering (Chiu 1994). When there are a few input variables, the grid partition is considered a suitable method for data classification. However, in this study because of many input variables and the requirement for considerable membership functions, the subtractive clustering method was utilized. For example, if we have 11 input variables and for each input variable three membership functions, the rules will be \( 3^{11} \) rules (177,147 rules) and the calculation of the parameters of this model will therefore be very complex (Noori et al. 2009). Therefore, in this study subtractive fuzzy clustering was used to establish the rule-based relationship between the input and output variables.

The subtractive clustering method assumes that each data point is a potential cluster centre and calculates a measure of the likelihood that each data point would define the cluster centre, based on the density of
surrounding data points. The algorithm selects the data point with the highest potential as the first cluster centre, then removes all data points in the vicinity of the first cluster centre, in order to determine the next data cluster and its centre location, and then iterates this process until all data is within the radii of a cluster centre (Yetilmezsoy et al. 2011). There are four algorithm parameters of subtractive clustering: range of influence, squash factor, accepted ratio and rejected ratio (Cakmakci et al. 2010; Yager, Filev 1994).

Subtractive clustering was developed by Chiu (1994) in order to estimate both the number and initial locations of cluster centres. Consider a set $T$ of $N$ data points in a $D$-dimensional hyper-space, where each data point $W_i$ ($i = 1, 2, \ldots, N$) $W_i = (x_i, y_i)$, where $x_i$ denotes the $p$ input variables and $y_i$ is the output variable. The potential value $P_i$ of the data point is calculated by Eq. (7):

$$ P_i = \sum_{j=1}^{N} e^{-\alpha ||W_i - W_j||^2}, $$

where $\alpha = 4/r^2$, $r$ is the radius defining a $W_i$ neighborhood, and $|| \cdot ||$ denotes the Euclidean distance (Wei et al. 2011).

The data point with many neighbouring data points is chosen as the first cluster centre. To generate the other cluster centres, the potential $P_i$ of each data points $W_i$ is revised by Eq. (8):

$$ P_i = P_i - P_k^* \exp\left(-\beta \| W_i - W_k^* \|^2 \right), $$

where $W_k^* = (x_k^*, y_k^*)$ is the location of the $k$th cluster center and $P_k^*$ is its potential value.

At the end of the clustering process, the method obtains $q$ cluster centres and $D$ corresponding spreads $S_i$, $i = (1, \ldots, D)$. Then we define their membership functions. The spread is calculated according to $\beta$ (Wei et al. 2011).

4. ANFIS models for predicting Australia’s domestic LCCs passenger demand

4.1. ANFIS process

As Figure 3 shows, the study was undertaken in three discrete phases. In the first phase an extensive literature review was undertaken to identify the extant knowledge on the determinants of LCC airline passenger demand. The requisite data was then sourced for the candidate input and output variables (Section 4.2). This data was subsequently normalized following the recommendations of Ghassemzadeh et al. (2013) and Mittal et al. (2012). The following step involved the data input. The input of the data included the input data and output data in the form of a data array (Chen et al. 2010: 1187). The final action at this stage involved defining and partitioning the universe of discourse for the input variables using the subtractive clustering method (Cakmakci 2007; Wei et al. 2011).

The next step involved is generating the fuzzy inference system (FIS) (Chen et al. 2010; Efendigil et al. 2009). The initialization of the fuzzy system was performed using the genfis 2 command, which specifies the structure and initial parameters of the FIS with the training data matrix, number of membership functions (MFs), and the membership types associated with each input (Patil et al. 2011). Generally, the coefficients for the MFs are initially selected by trial and error, and subsequently, fine-tuned using the hybrid learning algorithm (Gao, Ovaska 2002).

The FIS parameters from the training datasets were then optimised, using the least square method and the backpropogation gradient descent method for training the forecasting ANFIS models (Wei et al. 2011; Yetilmezsoy et al. 2011). The training of the study’s data was performed automatically in the ANFIS system and an array of training errors was obtained (Chen et al. 2010: 1187). Following training, an ANFIS model with a forecasting function was obtained for output forecasting (Bagheri et al. 2014; Chen et al. 2010: 1187). The models computed the overall output as a summation of all the incoming signals (Efendigil et al. 2009). Finally, a performance index, based on $R$, MAPE, MSE and RMSE (see Section 4.5 below), was established to evaluate the performance of the models.

4.2. Data sources

The availability of a consistent data set allows the use of quarterly data for the period 2002 to 2012. The data used
in the ANFIS models were sourced from a variety of sources. Data on Australia's real GDP and real GDP per capita, Australia's unemployment numbers, population size and recorded bed capacities at Australia's tourist accommodation establishments are from the Australia Bureau of Statistics (ABS). Australia's real interest rates are from the Reserve Bank of Australia (RBA). The airfare data are from the Bureau of Infrastructure, Transport and Regional Economics (BITRE) (airline yields are used as a proxy of the average airline fares and are based on Australia's real best discount air fares). The data on Australia's LCC domestic enplaned passengers and revenue passenger kilometres performed (RPKs) are from the Bureau of Infrastructure, Transport and Regional Economics (BITRE), Qantas Group, Tiger Airways and Virgin Australia reports and websites. 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, the consumer price index at 2011 constant prices was used (Ba-Fail et al. 2000).

Four dummy variables were included in the modelling. The first dummy variable explained the impact of the evolving Virgin Australia's business model from a low cost carrier model to a full service network carrier (FSNC) (Whyte et al. 2012) on Australia's low cost carrier traffic (enplaned passengers and RPKs). Australia's low cost carriers' traffic in Australia has decreased significantly since 2011, primarily due to this transition in Virgin Australia's business model evolution. Thus, the dummy variable reflecting the Virgin Australia's changing business model (DUMMY 1) is zero for the period from Quarter 1 2002 to Quarter 4 2010 and one from Quarter 1 2011 to Quarter 2 2012.

The second dummy variable accounted for the loss of capacity following the collapse of Ansett Australia. At the time of its collapse in 2001, Ansett Australia's domestic Australian market share was 35 per cent (Virgin Blue held around 10 per cent and Qantas had a 55 per cent market share) (Prideaux 2003). Ansett Australia experienced financial problems and was placed into receivership on September 14, 2001 (Easdown, Wilms 2002). The collapse of Ansett Australia had a major impact on the tourism industry, especially in regional areas where Ansett's subsidiaries provided substantial capacity. Whilst the other incumbent airlines increased seating capacity, the demand for seats exceeded supply for several months (Prideaux 2003).

The third dummy variable accounted for the impact of the Global Financial Crisis (GFC) during the period 2007 to 2009, whilst the fourth dummy variable accounted for the impact of the Commonwealth Games held in Melbourne from 15 to March 26, 2006.

In this study, each input/output pair contains 11 inputs (that is, GDP, air fare, population, unemployment, bed spaces, jet fuel prices, interest rates and 4 dummy variables and one output (RPKS or PAX). The input data are Australia's low cost carriers enplaned passengers and RPKs performed.

Prior to training the data in the ANFIS, it is important to process the data into patterns. Training and testing pattern vectors are formed. Each pattern is formed with an input condition vector as well as the corresponding target vector. The scale of the input and output data is an important matter for consideration, particularly when the operating ranges of process parameters are different. The normalizing of the data ensures that the ANFIS will be trained effectively, without any particular variable skewing the results significantly. Consequently, all the input parameters are of equal importance in training the ANN (Baseri 2011).

In this study all data were, therefore, normalized prior to use in the training phase using Equation (9). The data normalization was applied to transform the data to a symmetric distribution which improves the model performance since the data appear to more closely satisfy the assumptions of a statistical inference procedure also following the transformations of variables (Ghassemzadeh et al. 2013). The data is normalized using the following equation:

$$ x_{\text{norm}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} $$  

where $x_{\text{norm}}$ is the normalized value, $x$ is the actual value, $x_{\text{max}}$ is the maximum value, and $x_{\text{min}}$ is the minimum value (Kalkhaheh et al. 2012).

There are several advantages of normalizing data prior to processing in the ANFIS for prediction. One advantage is to avoid attributes in greater numeric ranges dominating those of smaller data ranges. The second advantage is to avoid numerical difficulties experienced during the calculation (Mittal et al. 2012). With data normalization, the data are scaled so they fall within a pre-specified range, such as $[0, 1]$ (Mitsa 2010). In this study's modelling process, all data values were scaled in the range between 0 and 1 using Equation (9). A further advantage of normalizing the data is that the normalization also removes any arbitrary effects of the similarity between objects whilst also increasing the answer rate data to the input signal (Mittal et al. 2012).

4.3. ANFIS models setup

In this study, the ANFIS (Adaptive Neuro-Fuzzy Inference System) Editor GUI (graphical user interface) in
the Fuzzy Logic Toolbox 2.2.16 within the framework of MATLAB R2012b (8.0.0.783) (The MathWorks, Inc., USA) software was used for modelling and simulation purposes.

The Sugeno ANFIS network setup process is conducted with 16 membership functions and the membership function type is Gaussian. The architecture of the study’s ANFIS is depicted in Figure 4. The hybrid learning algorithm was used for the ANFIS models.

The neuro-fuzzy models were run for each combination of model parameter with varying numbers of epochs to avoid the possible over-fitting of the models (Efendigil et al. 2009). The Gaussian-curve membership

Fig. 4. The optimum ANFIS model architecture for forecasting Australia’s LCCs enplaned passengers and RPKs.

Fig. 5. Initial and final Gaussian membership functions for the ANFIS models
function and 16 rules is the optimum architecture for the two ANFIS models. The generated membership functions are able to display the interactions and relationships between the various ANFIS levels. Figure 5 shows the fine curves of the trained models with smooth curve interaction for each parameter suggesting the best fit of the developed models (Mittal et al. 2012).

In this study, the ANFIS model was structured for forecasting Australia's low cost carrier air travel demand using the Sugeno approach with eleven inputs and one output. The "product" function is used for linking the rules together, "weighted average" is used for rule defuzzification, and the subtractive clustering algorithm partition method is applied to generate the optimum 16 fuzzy rule base sets (Efendigil et al. 2009), where the membership function's shape in input layer is set as the Gaussian membership function and the shape of the linear membership function is used in the output layer. Examples of 2 of the model's 16 rules are as follows:

Rule 1: If (fare is mf1) and (pop is mf1) and (gdp is mf1) and (unemp is mf1) and (fuel is mf1) and (int is mf1) and (accom is mf1) and (d1 is mf1) and (d2 is mf1) and (d3 is mf1) and (d4 is mf1) then pax is 0.48*fare + 0.06*pop + 0.10*gdp - 0.24*unemp - 0.12*fuel - 0.10*int - 0.18*accom + 2.3*10*d1 - 9.7*10*d2 - 8.3*15*d3 + 1.7*7*d4 + 0.0042.

Rule 16: If (fare is mf16) and (pop is mf16) and (gdp is mf16) and (unemp is mf16) and (fuel is mf16) and (int is mf16) and (accom is mf16) and (d1 is mf16) and (d2 is mf16) and (d3 is mf16) and (d4 is mf16) then pax is 0.08*fare + 0.10*pop + 0.10*gdp + 0.003*unemp + 0.13*fuel + 0.09*int + 0.15*accom + 2.1*10*1 + 1.9*18*2 + 1.8*29*d1 + 0.18.

4.4. Data training

Training is a key part of the ANFIS model development process. The training process is used to optimize the model, and the subsequent testing process is used to check the performance and, consequently, the generalization ability of the developed model (Mehta, Jain 2009). In this study, the testing data subset was independent from the training data set and was used to train the ANFIS model. The testing data set was utilised to verify the accuracy and effectiveness of the ANFIS model (Azadeh et al. 2010; Galavi, Shui 2012; Übeyli et al. 2010). The data was therefore divided into two randomly selected groups: the first group of 36 data was used as the training set (85% of the overall data), and the remaining group of 6 data was used for verifying and testing the robustness of the ANFIS-based prediction models (Yetilmezsoy et al. 2011).

The task of the learning algorithm for the study's ANFIS architecture is to tune all modifiable parameters, that is, \((a_1, b_1, c_1)\) and \((p_1, q_1, r_1)\), to ensure that the ANFIS output matches the training data. When the premise parameters \(a_1, b_1, c_1\) of the membership function are fixed, the output of the ANFIS can be expressed as (Übeyli et al. 2010):

\[
f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2.
\]

Substituting Equation (4) into 10 yields:

\[
f = \bar{w}_1 f_1 + \bar{w}_2 f_2.
\]

After further substitution of the fuzzy if-then rules into Equation (11), it becomes:

\[
f = \bar{w}_1 (p_1 x + q_1 y + r_1) + \bar{w}_2 (p_2 x + q_2 y + r_2).
\]

Following rearrangement, the output can be expressed as:

\[
f = \bar{w}_1 (p_1 x + q_1 y + r_1) + \bar{w}_2 (p_2 x + q_2 y + r_2)
\]

which is a linear combination of the modifiable consequent parameters: \(p_1, q_1, p_2, q_2, r_1\) and \(r_2\) (Übeyli et al. 2010: 682).

The least squares estimation (LSE) method can be utilised quite easily to identify the optimal values of these parameters (Übeyli et al. 2010). Normally a gradient based method is utilized for the ANFIS learning procedure. However, this method is known for its very slow performance and the tendency to become trapped in a local minimum (Kablan 2009). This study used a standard hybrid learning algorithm as proposed by Jang (1993), which utilises a combination of the steepest gradient and the least squares estimation (LSE) (Übeyli et al. 2010). Each epoch of this hybrid learning procedure compromises a forward pass and back propagation (Chen et al. 2010). In the forward pass, the functional signals proceed forward to layer 4 and the resulting parameters are identified by the least square estimate (Kablan 2009; Yan et al. 2010). Once the optimum consequent parameters are found, the backward pass immediately commences (Efendigil et al. 2009; Übeyli et al. 2010). In the backward pass, the error rates propagate backward and the premise parameters are updated by the gradient descent (Yan et al. 2010; Yilmaz, Kaynar 2011). The output of the ANFIS is calculated by employing the consequent parameters that are found in the forward pass. The output error is utilized to adapt the premise parameters by means of a standard backpropagation algorithm (Übeyli et al. 2010). It has been proven that this hybrid algorithm is highly efficient in ANFIS training (Jang 1993; Kablan 2009; Übeyli et al. 2010). Table 1 presents a summary of the learning methods used in the training phase.

Table 1. Summary of the study's ANFIS training algorithm

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Forward pass</th>
<th>Backward pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premise</td>
<td>Fixed</td>
<td>Gradient descent</td>
</tr>
<tr>
<td>Consequent</td>
<td>LSE</td>
<td>Fixed</td>
</tr>
<tr>
<td>Signals</td>
<td>Node outputs</td>
<td>Error rates</td>
</tr>
</tbody>
</table>
Each ANFIS model used 36 training data in 1–400 training epochs (Übeyli et al. 2010). Figure 6 shows the training curve of the ANFIS (PAX Model) with a root mean square error (RMSE) of 0.000000378. Figure 7 shows the training curve of the ANFIS RPKs model with an RMSE of 0.000000376. These figures display the level of modelling accuracy in terms of the error achieved (Mittal et al. 2012).

A comparison between the actual and the ANFIS predicted PAX and RPKs models’ values following the completion of the training are presented in Figures 8 and 9, respectively. The two figures show that the ANFIS system is well-trained to model Australia’s actual low cost carrier passenger demand, as measured by both passengers carried and revenue passenger kilometres performed.

4.5. Model evaluation goodness of fit measures

Goodness-of-fit (GOF) statistics are useful when comparing results across multiple studies, for examining competing models in a single study, and also for providing feedback on the level of knowledge about the uncertainty involved in the phenomenon of interest (Kunt et al. 2011). Five measures were used in the present study: the coefficient of determination ($R^2$), the mean absolute error (MAE), the root mean square error (RMSE), the mean square error (MSE) (Yetilmезsoy et al. 2011) and the mean absolute percentage error (MAPE) (Azadeh et al. 2010; Chen et al. 2010).

For evaluating the ANFIS models, the root mean squared error (RMSE), the mean absolute error (MAE), the mean absolute percentage error (MAPE), the mean square error (MSE), and the coefficient of determination ($R^2$), were calculated using Equations (14)–(17):

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

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{t_i - td_i}{t_i} \right|;$$

$$\text{MAPE} = \frac{1}{N} \left( \sum_{i=1}^{N} \left| \frac{t_i - td_i}{t_i} \right| \right) \times 100;$$

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

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (t_i - td_i)^2}{\sum_{i=1}^{N} (t_i - \bar{t})^2},$$

where $t_i$ is the actual values; $td_i$ is the predicted values; $N$ is the total number of data (Tiryaki, Aydın 2014: 104).

4.6. ANFIS modelling results

The computation of the data for the ANFIS models was conducted using the software Matlab. The ANFIS training algorithms, including the gradient method and the least squares estimation method, were embedded in the software of Matlab’s fuzzy inference toolbox. The main computation procedure involved four steps. The first step is the data input. The input of the model data includes the input data and the output data in the form of a data array. The second step is generating the fuzzy inference system. The third step is utilising the ANFIS training function in the toolbox for the training of the input data. The training of the data will be performed automatically in the system and an array of training error will be obtained. Following training, an ANFIS model with the forecasting function was obtained for the output forecasting as the last step (Chen et al. 2010). Figure 10 depicts the Australia’s low cost carrier passengers/RPKs demand forecasting system according to the Sugeno approach.
The ANFIS was trained using Matlab 7.0 with the various possible combinations of the subtractive clustering parameters (range of influence (ROI) = 0.45–0.60, squash factor (SF) = 1.20–1.35, accept ratio (AR) = 0.40–0.55 and reject ratio (RR) = 0.10–0.20) for the range of the epoch number from 1–400 epochs. The constructed ANFIS model was manipulated until the best settings were obtained, based on the lowest RMSE value. The hybrid learning algorithm was applied in the training phase. The data are normalized to the scale [0,1] in order to increase the training performance. The training process stopped whenever the maximum epoch number was reached or the training error goal was achieved.

The root mean square errors (RMSE) became steady after running 20 epochs of PAX and RPKs training data. The final convergence values were 0.00000378 and 0.00000376 for the PAX and RPKs models, respectively.

The parameter in the subtractive clustering fuzzy inference system comprises the range of influence (ROI), squash factor (SF), accept ratio (AR) and reject ratio (RR) (Yetilmzesoy et al. 2011). The constructed ANFIS models were manipulated by changing the parameters of clustering systematically around their default values until the best settings were obtained based on the lowest RMSE value. It is found that the optimum ANFIS structure of the PAX model with ROI = 0.52, SF = 1.25, AR = 0.50 and RR = 0.15 returns the lowest value of the RMSE at 0.00000378 and the RPKs model with ROI = 0.53, SF = 1.25, AR = 0.50 and RR = 0.15 returns the lowest value of the RMSE at 0.00000376. The optimum ANFIS model architecture for the forecasting of the LCCs enplaned passengers and RPKs is shown in Figure 10.

Following training, the ANFIS model for forecasting Australia’s LCCs enplaned passengers and RPKs was validated by selecting six data points, which are different from the other 36 points used for the ANFIS training (Al-Ghandoor et al. 2012). Each validation data point was fed into the system and then the Australia’s predicted LCCs enplaned passengers and RPKs values were computed and compared to the actual values. The performance index of training, testing and overall data of PAX and RPK model was calculated, as shown in Table 2.

Table 2 shows that both the PAX and RPKs ANFIS models achieve a very satisfactory predictive accuracy. Both models show that the MAE, MAPE, MSE and RMSE are very low for the training, testing and overall data sets.

Table 2. Performance index of the ANFIS models for the training, testing and overall data set

<table>
<thead>
<tr>
<th>Performance index</th>
<th>PAX model</th>
<th>RPK Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train data</td>
<td>Test data</td>
</tr>
<tr>
<td>MAE</td>
<td>0.001</td>
<td>0.047</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.38%</td>
<td>8.38%</td>
</tr>
<tr>
<td>MSE</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.001</td>
<td>0.058</td>
</tr>
</tbody>
</table>

The overall estimated and the actual value of LCC enplaned passengers and RPKs were regressed and, as Figure 11 shows, the $R^2$ are very high, being around 0.9949 and 0.9953 for the PAX and RPKs models, respectively.

All actual and predicted values of Australia’s LCCs enplaned passengers (PAX) and RPKs models are plotted in Figures 12 and 13, respectively. These figures clearly show the fit of the ANFIS to the actual data, indicating the extremely high estimation accuracy of the study’s ANFIS models.

Fig. 11. Comparison of the estimated and the actual values of the ANFIS model for forecasting Australia’s LCCs enplaned passengers and RPKs.
5. Conclusions

This study has proposed and tested two ANFIS models for predicting Australia’s domestic low cost carriers’ demand, as measured by enplaned passengers and revenue passenger kilometres performed (RPKs). Sugeno fuzzy rules were used in the ANFIS structure and the Gaussian membership function and linear membership functions were also developed. The hybrid learning algorithm and the subtractive clustering partition method were used to generate the optimum ANFIS models. The data was normalized to the scale [0,1] in order to increase the model’s training performance. The results found that the mean absolute percentage error (MAPE) for the overall data set of LCCs enplaned passengers (PAX) and RPKs models were 1.52% and 1.17%, respectively.

It can be concluded that the ANFIS is an approach that can be used to model and predict Australian low cost carrier passenger air travel demand effectively. The originality of this study is the use of the adaptive neuro-fuzzy inference system (ANFIS) approach which has not been previously used to forecast Australia’s low cost carrier passenger air travel demand. The ANFIS models produced very satisfactory results and showed high forecasting accuracy. The application of the ANFIS approach for the prediction of other air transport demand may also be worthy of future research and interest.

References


Cakmakci, M.; Kinaci, M.; Bayramoglu, Y. et al. 2010. A modelling approach for iron concentration in sand filtration


http://dx.doi.org/10.3846/16484142.2011.635465


