

# A Forecasting Tool for Predicting Australia's Domestic Airline Passenger Demand Using a Genetic Algorithm

Panarat Srisaeng<sup>1</sup>, Glenn Baxter<sup>2</sup>, Steven Richardson<sup>3</sup>, Graham Wild<sup>2</sup>

**ABSTRACT:** This study has proposed and empirically tested for the first time genetic algorithm optimization models for modelling Australia's domestic airline passenger demand, as measured by enplaned passengers (GAPAXDE model) and revenue passenger kilometres performed (GARPKSDE model). Data was divided into training and testing datasets; 74 training datasets were used to estimate the weighting factors of the genetic algorithm models and 13 out-of-sample datasets were used for testing the robustness of the genetic algorithm models. The genetic algorithm parameters used in this study comprised population size ( $n$ ): 200; the generation number: 1,000; and mutation rate: 0.01. The modelling results have shown that both the quadratic GAPAXDE and GARPKSDE models are more accurate, reliable, and have greater predictive capability as compared to the linear models. The mean absolute percentage error in the out of sample testing dataset for the GAPAXDE and GARPKSDE quadratic models are 2.55 and 2.23%, respectively.

**KEYWORDS:** Australia, Forecasting method, Genetic algorithm, Domestic airlines, Air transport.

## INTRODUCTION

Australia's airline industry was born on connecting regional communities to the country's major cities (Baker and Donnet 2012). Due to the vast distances across the country as well as between urban centres, Australia is heavily reliant upon its air transport industry (Nolan 1996). Australia's air transport industry was historically tightly controlled by the government. However, 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 low-cost carriers now have around 35% market share, with the 2 major incumbent LCCs being Jetstar and Tiger Airways. Qantas and Virgin Australia are the two present incumbent full service network carriers (FSNCs) (Srisaeng *et al.* 2014).

Reliable forecasts of air transport activity play a vital role in the planning processes of States, airports, airlines, engine and airframe manufacturers, suppliers, air navigation service providers, and other relevant bodies. In addition to assist States in facilitating the orderly development of civil aviation and to aid all levels of government in the planning of air space and airport infrastructure — for example, air traffic control (ATC), airport air side and landside facilities —, reliable forecasts also assist aircraft manufacturers in planning future aircraft types (in terms of size and range) and when to develop them (International Civil Aviation Organization 2006).

Forecasting passenger transport demand is viewed as being of critical importance for airlines as well as for investors since investment efficiency is greatly influenced by the accuracy and

**1.** Suan Dusit University - School of Tourism and Hospitality Management - Bangkok - Thailand **2.** RMIT University - School of Aerospace, Mechanical, and Manufacturing Engineering - Melbourne/VIC - Australia. **3.** Edith Cowan University - School of Engineering - Perth/WA - Australia.

**Author for correspondence:** Graham Wild | RMIT University - School of Aerospace, Mechanical, and Manufacturing Engineering | 124 La Trobe St | Melbourne/VIC 3000 - Australia | Email: graham.wild@rmit.edu.au

**Received:** 02/26/2015 | **Accepted:** 10/28/2015

adequacy of the estimation performed (Blinova 2007). Air traffic forecasts are therefore one of the key inputs into an airline's fleet planning, 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 (Ba-Fail *et al.* 2000). Forecasts also assist airlines in their decision-making regarding the development of infrastructure facilities, thereby enhancing services provided to their passengers (Abed *et al.* 2001). While all of these activities have equivalents in other industries, airlines suffer from much narrower profit margins. In 2014 the average global margin was 3.2% (International Air Transport Association 2015). This makes the predictive capability of forecasts much more significant in the aviation industry. It also makes the use of advanced and innovative models a very interesting prospective for improving all short-term to midterm planning activities.

Despite the significance of Australia's domestic airline market sector, there has been no previously reported study that has developed and empirically tested genetic algorithm-based models for forecasting Australia's domestic airline passenger demand. The primary objective of this study is to address this apparent research gap in the literature. In order to address the research objective, various forms of mathematical expressions were proposed and tested. The study also sought to examine whether the genetic algorithm approach is a useful tool for this application. Genetic algorithm enplaned passengers (GAPAXDE) and genetic algorithm revenue passenger kilometres performed (GARPKSDE) models are proposed to forecast Australia's domestic airline quarterly enplaned passengers and revenue passenger kilometres performed, respectively. Airline passenger traffic can be measured by the number of passengers carried (enplaned passengers) and also by revenue passenger kilometres performed (Belobaba 2009; Holloway 2008).

## TRADITIONAL AIR TRAVEL DEMAND FORECASTING APPROACHES

In the air transport industry, many service providers and government regulatory agencies follow the International Civil Aviation Organization (ICAO) *Manual on Air Traffic Forecasting*. This manual was originally developed in 1985 using traditional modelling techniques (Alekseev and Seixas 2009). Historically, multiple linear regression (MLR) models have generally been used to forecast airline passenger traffic

demand (see, for example, Aderamo 2010; Ba-Fail *et al.* 2000; Bhadra 2003; Kopsch 2012; Sivrikaya and Tunç 2013).

Genetic algorithms (GAs) are an alternative artificial intelligence-based forecasting approach that could potentially be used for forecasting air travel demand. GAs are powerful stochastic search techniques that are based on the principle of natural evolution (Kunt *et al.* 2011). The theoretical basis for the GAs is the "Schema Theorem" (Holland 1975). This states that individual chromosomes with good, short, low-order schemata or building blocks (that is, beneficial parts of the chromosome) receive an exponentially increasing number of trials in successive generations (Hurley *et al.* 1998). GAs differ substantially from traditional optimization methods because they search using a population of points in parallel rather than a single point in order to obtain the best solution (Akgüngör and Doğan 2009). Furthermore, GAs allow for a broader and more global search of the solution space. Indeed, the aim of GAs is to determine the optimal solution to a given problem under study (Carvajal-Rodríguez and Carvajal-Rodríguez 2009).

The GAs forecasting approach has been applied to a wide range of disciplines in recent times, including electric energy estimation (Ozturk *et al.* 2005), energy demand prediction (Ghanbari *et al.* 2013), housing price forecasting (Jirong *et al.* 2011), tourism demand forecasting (Hernández-López and Cáceres-Hernández 2007; Hong *et al.* 2011), traffic accident severity prediction (Akgüngör and Doğan 2009; Kunt *et al.* 2011), and transport energy demand prediction (Haldenbilen and Ceylan 2005). Despite the reported benefits of GAs, there has only been one reported study that has applied GAs in the aviation industry. Sineglazov *et al.* (2013) have proposed a GAs forecasting method for solving the problems of forecasting experienced in the aviation industry. These authors have also noted that their GAs may be applicable for forecasting regional aviation facilities and other industrial sectors that have demand patterns similar to those experienced by airlines.

## GENETIC ALGORITHMS: A BRIEF OVERVIEW

GAs are based on the genetic process of biological organisms that are explained by the principles of natural selection and survival of the fittest (Akgüngör and Doğan 2009). GAs are therefore similar to the natural evolution process where the population of a given species adapts to a natural environment, a population of designs is subsequently created and then permitted to evolve in order to adapt to the design environment that is being considered (Azadeh *et al.* 2011). GAs encode a

possible solution to a specific problem on simple chromosome string like data structure and apply specific operators to these structures so as to preserve important information (Jones and Romil 2004).

The principal strength of GAs is their adaptive and self-organizing capabilities. The basic operations of GAs include selection, a crossover of genetic information between reproducing parents and a mutation of genetic information which affects the binary strings characteristic in natural evolution (Ozturk *et al.* 2005). If GAs are suitably encoded, then they can be used to solve real-world problems by mimicking this process (Akgüngör and Doğan 2009). In general GAs are a meta-heuristic algorithm, which can be used to determine a close to optimal solution to a specific problem. There are other evolutionary algorithms and other meta-heuristics approaches. The choice of the GA is based on the specific strengths previously identified, especially when dealing with large solutions spaces, based on large combinations of a relatively small number of inputs.

#### Genetic Algorithm Process

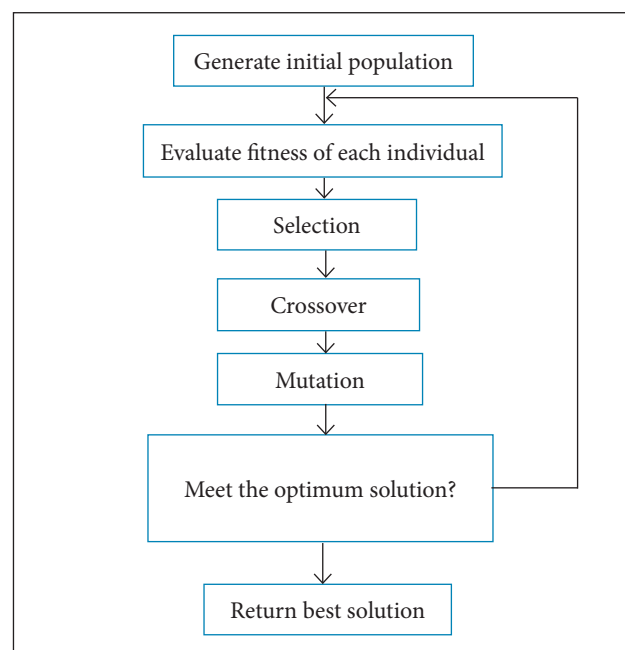
The GA commences with a population of solutions (chromosomes), which is termed “population”, represented by coded strings (typically 0 and 1 binary bits) as the underlying parameter set of the optimization problem (Kunt *et al.* 2011). Each individual in the population is called a chromosome, and these represent the candidate solution to the problem at hand (Gen and Cheng 1997). GAs generate successively improved populations of solutions (better generations) by applying three main genetic operators: selection, crossover, and mutation (Amjadi *et al.* 2010; Coelho *et al.* 2014; Kunt *et al.* 2011).

With a GA it is a requirement to create an initial population to serve as the starting point. This population can be created randomly or by using specialized, problem-specific information on the specific problem being investigated (Godinho and Silva 2014). Over a wide range of applications, an initial population size of between 30 and 100 has often been used (Goldberg 1989). Chromosomes evolve through successive iterations, which are termed “generations” (Rotshtein and Rakytyanska 2012). During each generation the chromosomes are evaluated, using some measures of fitness (Ozturk *et al.* 2005). To create the following generation, new chromosomes, called off-spring, are formed by (i) merging two chromosomes from a current generation using a crossover operator, or (ii) by modifying a chromosome using a mutation operator. A new generation is formed by (i) selecting, according to fitness values, some of the parents and the off-spring

whilst and (ii) rejecting others to keep the population size constant. Fitter chromosomes have a higher probability of being selected. Following several generations, the algorithms converge to a good population, which should contain the optimal or sub-optimal (close to optimal) solution to the problem at hand (Gen and Cheng 1997).

The GA works with operations that are performed based on fitness evaluation. The fitness indicates the goodness of design, and, accordingly, the objective function is a logical choice for the fitness measure (Ozturk *et al.* 2005). Fitness evaluation involves defining an objective or fitness function against which each chromosome is tested for suitability for the environment that is being considered in the study (Hurley *et al.* 1998). The GA selects the fittest members of the population based upon the best fitness value (Haldenbilen and Ceylan 2005). The present study’s GA process is illustrated in Fig. 1.

GAs work according to selection rules as defined by the laws of evolutionary genetics (Ozturk *et al.* 2005). The selection function chooses parents for the next generation based on their scaled values from the fitness scaling function in which the stochastic uniform selection function was used (Kunt *et al.* 2011). The selection mechanism consists of algorithms that mimic natural selection and select the best combination from a set of competing solutions. These selection algorithms (for example, rankings) yield preferences for the best performers (Hu 2002).



**Figure 1.** The study’s genetic algorithm process (adapted from Amjadi *et al.* 2010).

When using a GA, it is a requirement to select chromosomes from the current population for reproduction. The selection procedure selects two parent chromosomes based on their fitness values, where the better the fitness value, the higher the probability that a chromosome is selected by the GA. The parent chromosomes are subsequently used by the crossover and mutation operators to produce two offspring for the new population. This selection/crossover/mutation cycle is repeated until the new population contains  $2n$  chromosomes. This means the process stops after  $n$  cycles (Hurley *et al.* 1998).

Crossover is achieved by exchanging coding bits between two mated strings in the GA (Kunt *et al.* 2011). Once a pair of chromosomes has been selected, crossover can then occur in order to produce offspring (Cervantes and Costillo 2013). This operation is executed by selecting two mating parents, randomly selecting two sites on each of the chromosomal strings and subsequently swapping the strings between the sites among the pair (Ozturk *et al.* 2005). Thus, parents produce offspring having different genetic structures that include some mix of their chromosomes set (Akgüngör and Doğan 2009). An illustration of the crossover operation is as follows (Ozturk *et al.* 2005):

Parent 1 = 1010101011  
 Parent 2 = 1001000111  
 Child 1 = 1010000111  
 Child 2 = 1001101011

The crossover process is repeated from one generation to another until one individual dominates the population or until the predetermined numbers of generations are reached. Conversely, crossover is not normally applied to all pairs of individuals selected for mating (Akgüngör and Doğan 2009). The crossover operation is carried out with a probability  $p_c$ . Typical probability values range from 0.2 to 0.8 (Ozturk *et al.* 2005).

The mutation operation serves a critical role in GAs either through the replacement of genes lost from the population during the selection process or by providing genes that were not included in the initial population (Akgüngör and Doğan 2009). In GAs, the mutation operator is invoked with a low probability ( $p_m$ ) at a randomly selected site on chromosomal string of the randomly chosen design. The operation consists in switching between 0 – 1 or vice versa (Ozturk *et al.* 2005). Mutation is therefore randomly applied with a small probability,

which is typically in the range between 0.001 and 0.01 and modifies genes in the chromosomes. The effect of mutation on a binary string is illustrated as follows (Akgüngör and Doğan 2009):

Offspring                      10101110 11  
 Mutated offspring            10101110 01

The difference between the offspring and the mutated offspring is underlined; this is the mutated gene (a “1” has been randomly switched to a “0”).

## GA MODELS FOR FORECASTING AUSTRALIA'S DOMESTIC AIRLINE PASSENGER DEMAND

### THE GAPAXDE AND GARPKSDE DATA AND VARIABLES SELECTION

In this study, GAPAXDE and GARPKSDE genetic algorithm prediction models have been proposed. It is important to note that the factors that influence air travel demand are complex (Doganis 2009; Vasigh *et al.* 2008). Each factor is composed of elements that can stimulate or 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). During the models development process and based on an extensive literature review of the factors that influence air travel demand, 11 variables were considered for inclusion and testing as independent variables in the two GA models:

- $X_1$  Australia's real best discount economy airfare;
- $X_2$  Australia's population size;
- $X_3$  Australia's real Gross Domestic Product (GDP);
- $X_4$  Australia's real GDP *per capita*;
- $X_5$  Australia's unemployment size;
- $X_6$  Australia's real interest rates;
- $X_7$  World jet fuel prices;
- $X_8$  Recorded bed capacities at Australia's tourist accommodation establishments;
- $X_9$  Ansett Australia's collapse in 2001 (Dummy 1);
- $X_{10}$  Sydney Olympic Games in 2000 (Dummy 2); and
- $X_{11}$  Melbourne Commonwealth Games in 2006 (Dummy 3).

Each of these factors is discussed in order to provide an insight into how they influence air travel demand.

The principal driver of air travel demand is economic growth (Wensveen 2011). The most important socio-economic variable affecting the demand for leisure travel is personal or household income, since leisure trips are paid for by the passenger, who may also be paying for a spouse and one or more children (Doganis 2009). Furthermore, higher levels of economic activity will lead to greater demand for air transport services, because of the increased business requirements and generally higher spending of consumers (Bureau of Transport and Communication Economics 1994). During periods of economic growth (and when consumer confidence is strong), air travel demand grows. Conversely, when economies fall into recession or experience downturns, unemployment grows, and consumer confidence declines, individuals often postpone discretionary travel and other luxury purchases (Dempsey and Gesell 1997). Real GDP and real GDP *per capita* were used in this study to measure the effect of income on Australia's domestic air travel demand (enplaned passengers and revenue passenger kilometres performed).

The impact of Australia's demographic changes was considered through Australia's population and the number of unemployed people (Tsekeris 2009; Young and Wells 2011). 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). Unemployment rates have also been reported as being a determinant of air travel demand (Clark *et al.* 2009; Wensveen 2011). *Ceteris paribus*, increasing levels of employment tend to increase air travel demand. Conversely, unemployment tends to depress air travel demand (McKnight 2010).

A decrease in the real cost of air travel also positively influences air traffic growth (Hanlon 2007; Holloway 2008). However, the measurement of the price of air travel is normally complicated by the presence of different air fare classes offered by airlines (International Civil Aviation Organization 2006). Hence, airline passenger yields are often used as a proxy for air fares, which can be difficult to obtain given the wide use of a variety of (and fluctuating number of) discount fares. When all other influencing factors remain constant, falling yields will tend to increase traffic volumes. Conversely, rising yields will tend to reduce traffic volumes, subjected to demand elasticities (Doganis 2009).

Air transport and tourism are inter-related (Bieger and Wittmer 2006). Consequently, there is a strong association between the levels of air travel demand and tourism demand. Air transport can influence tourism demand via a number of channels, with price being one of these key factors (Koo *et al.* 2013). The tourism industry makes a substantial and important contribution to the overall level of economic activity and employment in Australia. As is the case with most developed economies, Australia's tourism industry is heavily oriented towards domestic expenditure by Australian residents (Hooper and van Zyl 2011). Domestic tourism accounts for three-quarters of tourism expenditure in Australia with the balance accounted for by international tourist spending (Hooper and van Zyl 2011; Organization for Economic Cooperation and Development 2014). Thus, another explanatory variable included in the study relates to tourism attractiveness, which is expressed in terms of the tourist accommodation infrastructure, that is, the reported bed capacity (Tsekeris 2009).

There is also a variety of other factors that may influence air travel demand. For example, jet fuel prices (Gesell 1993) and real interest rates. Sharp increases in world oil prices have had important impacts on world air travel demand. In addition to the adverse impact on the global economy, airlines are often forced to increase air fares to cover the higher fuel costs, which often have a detrimental impact on air travel demand. This is because increases in oil prices result in higher air fares and therefore make leisure travel more expensive (Li 2010). Furthermore, interest rates influence the balance between expenditure and saving (Cook 2007). High interest rates will inhibit economic activity, which can have a dampening effect on airline traffic (Wensveen 2011).

Three dummy variables were also included in the GA models. The first 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% (Virgin Blue held around 10% and Qantas had a 55% market share) (Prideaux 2003). Ansett Australia experienced financial problems and was placed into receivership on September 14, 2001 (Easdown and 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 dummy variable reflecting the collapse of Ansett Australia is

equal 1 for the period from Quarter 3 2001 to Quarter 2 2002 and 0 otherwise.

The second dummy variable 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 September 15, 2000). The Paralympics were also staged in Sydney over a 12-day period shortly after the conclusion of the Olympic Games (Madden 2002). During the Sydney Olympic Games, around 2,300 extra domestic flights were operated (Hensher and Brewer 2002) to satisfy extra passenger demand. Thus, the dummy variable reflecting the influence of the Olympic Games is equal 1 for Quarter 3 2000 and 0 otherwise.

The third dummy variable accounted for the impact of the Commonwealth Games held in Melbourne from March 15 to March 26, 2006. The 2006 Melbourne Commonwealth Games was the largest sporting and community event held in Victoria's history. The Commonwealth Games provided substantial economic benefits for the State of Victoria as well as for the tourism and airline industry (KPMG 2006). Therefore, the dummy variable reflecting the impact of the Commonwealth Games is equal 1 for Quarter 1 2006 and 0 otherwise.

The availability of a consistent data set allows the use of quarterly data for the period Quarter 4 1992 to Quarter 2 2014. This is the extent to which the full variables data sets are available. The data does not exist prior to 1992 for Australia's air fares (the Bureau of Infrastructure, Transport and regional Economics [BITRE] commenced publishing Australia's air fares time series data in October 1992). The data used in the GA models were sourced from a variety of sources. Data on Australia's real GDP, Australia's real GDP *per capita*, Australia's population size, Australia's unemployment numbers, and recorded bed capacities at Australia's tourist accommodation establishments (based on Australian tourist accommodation establishments with 15 rooms or more) 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 BITRE (airline yields are used as a proxy of average airline fares and are based on Australia's real best economy discount air fares). The data on Australia's domestic enplaned passengers and revenue passenger kilometres performed (RPKs) are from the BITRE as well. 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 (International Civil Aviation Organization 2006). The final point

to note about the variable selection and modelling is regarding the "non-stationary" nature of the economic variables considered above. First it should be noted that International Civil Aviation Organization (2006), the international authority on aviation, recommends the use of linear regression for forecasting air travel demand. However, it was also found that, when the data set in question was converted to "stationary" data (using both the first and second difference method), all models considered resulted in less predictive capability, then the models using the "non-stationary" data, in fact, were not even statistically significant (Srisaeng 2015).

## THE GAPAXDE AND GARPKSDE GENETIC ALGORITHM PROCESS

The GA utilised in this work was designed specifically for this application and has not been adapted from a previous GA. The goal is to determine an optimal (or close to optimal) subset of  $k$  independent variables (chosen from a set of  $n$  variables  $\{X_i; i = 1, 2, \dots, n\}$ ) which collectively provide the best predictive model of a dependent variable. Three models are considered to facilitate the investigation of as many relationships between the independent and dependent variables. These are:

$$\hat{y} = \sum_{i=1}^n w_i a_i X_i \quad (1)$$

$$\hat{y} = \sum_{\substack{i=1 \\ i \in l}}^n w_i a_i \ln(X_i) + \sum_{\substack{i=1 \\ i \notin l}}^n w_i a_i X_i \quad (2)$$

$$\hat{y} = \sum_{i=1}^n w_i a_i X_i + \sum_{i=1}^n \sum_{j=1}^i w_{ij} a_{ij} X_i X_j \quad (3)$$

where:

$i$  is the set of variables which only take positive values (that is, not zero or negative), as is  $j$  in Eq. 3.

Equation 1 captures simple linear relationships, Eq. 2 captures logarithmic relationships (common with socioeconomic variables) by allowing a variable to be described as either a linear or a log (but not both in the same solution), and Eq. 3 captures polynomial relationships and potential cross multiple relationships between the variables.

The coefficient component  $w_i$  indicates whether the variable  $X_i$  is included in the model, where  $w_i = 1$  if  $X_i$  is included, and  $w_i = 0$  if  $X_i$  is not included. Similarly,  $w_{ij}$  indicates whether the variable product  $X_i X_j$  is included in the quadratic model. In

what follows we will denote the vector of weight values as  $\mathbf{w}$  and define the set of feasible weight vectors  $W_k = \{\mathbf{w}: \mathbf{w} \cdot \mathbf{1} = k\}$ . That is,  $\mathbf{w}$  represents the switching “on” and “off” of genes, which, in this study, are variables in a multiple regression model.

Once  $\mathbf{w}$  is specified, the values of  $a_i$   $\{a_i: i = 1, 2, \dots, n\}$  (also  $\{a_{ij}: i = 1, 2, \dots, n; j = 1, 2, \dots, i\}$  for the quadratic model) are chosen to minimise the squared difference between the observed values of the dependent variable over  $m$  observations  $\{y_k: k = 1, 2, \dots, m\}$  and the corresponding predicted values (i.e. least squares). That is, denoting the vector of model coefficients as  $\mathbf{a}$ , we minimise:

$$LS(\mathbf{a}|\mathbf{w}) = \sum_{k=1}^m (y_k - \hat{y}_k)^2 \quad (4)$$

where:

$k$  is the integer representing the increments in the dependent variable, in this work, expressed as quarterly data.

### Objective Function

The goal is to determine the weight vector  $\mathbf{w}^*$  such that:

$$\mathbf{w}^* = \arg \min_{\mathbf{w} \in W_k} \left[ \min_{\mathbf{a}} LS(\mathbf{a}|\mathbf{w}) \right] \quad (5)$$

### Genetic Algorithm

If the number of independent variables ( $n$ ) is large, then the number of variable combinations of size  $K$  will also be large. Specifically, the number of combinations will be  $\binom{n}{k}$  for the linear model and  $(n^2 + 3n/2K)$  for the quadratic one. This may make it prohibitive to exhaustively evaluate all models with  $K$  independent variables. As such, the chosen approach to determine a close to optimal set of independent variables is to utilise a GA.

The stages of the GA implemented are outlined as follows:

- 1 Generate an initial population: An initial population ( $P_0$ ) is generated by randomly selecting a set of  $M$  solutions from the feasible solution set  $W_K$ . For each member of the initial population, we define the measure of fitness  $F(\mathbf{w})$  to be:

$$F(\mathbf{w}) = \min_{\mathbf{a}} LS(\mathbf{a}|\mathbf{w}) \quad (6)$$

- 2 Breed new population members for next generation: A pre-specified number  $B$  of new population members are bred at each generation. To breed each new population

member, we first choose two distinct parents from the existing population  $P_{i-1}$ , where  $i$  is now the generation number, with probabilities weighted by the inverse of the fitness measure  $F$  (i.e. lower values of  $F$  are associated with better solutions). That is, the probability of choosing solution for breeding each new solution is given by:

$$\Pr(\mathbf{w}) = \frac{\frac{1}{F(\mathbf{w})}}{\sum_{\mathbf{w}_j \in P_{i-1}} \frac{1}{F(\mathbf{w}_j)}} \quad (7)$$

Once the parents  $\mathbf{w}_1$  and  $\mathbf{w}_2$  are chosen, the child solution  $\mathbf{w}_c$  is bred using the following rules:

- (i) If  $\mathbf{w}_1(i) = \mathbf{w}_2(i) = 1$ , then  $\mathbf{w}_c(i) = 1$  (i.e. any variable/variable product that exists in both parent solutions is passed on to the child solution).
- (ii) The remaining variables/variable products in the child solution (i.e. to make up a total of  $K$ ) are randomly chosen from those where  $\mathbf{w}_1(i) = 1, \mathbf{w}_2(i) = 0$  or  $\mathbf{w}_1(i) = 0, \mathbf{w}_2(i) = 1$ .
- (iii) With probability  $P_{mut}$  (user specified), a breeding mutation occurs in which a randomly chosen variable/variable product that does not exist in either parent will exist in the child.

- 3 Introduce new migrating population members: At each new generation, a set of  $G$  new population members migrate into the population. These new population members are generated randomly in the same way as the members of the initial population.
- 4 Eliminate existing population members: In order to maintain a constant population size, a total of  $B + G$  members of the existing population must be discarded. Members are chosen with probabilities weighted by the fitness measure  $F$  (that is, higher values of  $F$  are associated with worse solutions), although the best solution is protected from elimination. The probability of choosing solution for elimination is given by:

$$\Pr(\mathbf{w}) = \frac{F(\mathbf{w})}{\sum_{\mathbf{w}_j \in P_{i-1}} F(\mathbf{w}_j)} \quad (8)$$

- 5 Form new population: The next population  $P_i$  is formed by combining the remaining (non-eliminated) population members from  $P_{i-1}$  with the new solutions that were bred and migrated into the population. Steps 2 to 5 are repeated for a predefined number of cycles,

or until a pre-specified number of generations pass without improvement.

### 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). Four measures were used in the present study: mean absolute error (MAE), the root mean square error (RMSE), mean square error (MSE) (Yetilmezsoy *et al.* 2011) and mean absolute percentage error (MAPE) (Azadeh *et al.* 2010; Chen *et al.* 2010).

For evaluating the GA models, RMSE, MAE, MAPE, and MSE were calculated using Eqs. 9 – 12:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (t_i - td_i)^2} \quad (9)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N \left| \frac{t_i - td_i}{t_i} \right| \quad (10)$$

$$MAPE = \frac{1}{N} \left( \sum_{i=1}^N \left| \frac{t_i - td_i}{t_i} \right| \right) \times 100 \quad (11)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (t_i - td_i)^2 \quad (12)$$

where:

$t_i$  are the actual values;  $td_i$  are the predicted values;  $N$  is the total number of data (Tiryaki and Aydın 2014).

## GAPAXDE AND GARPXSDE MODELLING RESULTS

To estimate model parameters, data was divided into two sets: training and testing datasets. The training dataset was used to estimate the weighting factors of the GA models, and testing dataset was saved for the testing purpose. The testing procedure is applied to obtain minimum relative error between estimated and actual values (Azadeh *et al.* 2007). In this study, the first group of 74 data was used as the training, and the remaining 13 out-of-sample data were used for verifying and testing the robustness of the GA models.

To identify the best fitness, the GA parameters were chosen as follows:

- Population size ( $n$ ): 200.
- Iterations (the generation number): 1,000.
- Mutation rate: 0.01.
- Probability of crossover: 1.

Prior to reviewing the modelling results, it is important to note that the GA modelling tested both the predictive capability of real GDP *per capita* and real GDP and Australia's population growth on Australia's domestic air travel demand in separate models. The modelling results showed that the inclusion of both real GDP and Australia's population size provided more robust and accurate model forecasting capability; that is, the models utilising real GDP and population gave better errors. After applying the GAPAXDE and GARPXSDE model procedures, the equations for forecasting Australia's domestic airline enplaned passengers (PAX) and RPKs were obtained based on the minimum sum of square errors between the observed and estimated data. These equations are 13 and 14, respectively.

The GAPAXDE<sub>lin</sub> and GARPXSDE<sub>lin</sub> present the linear models for forecasting Australia's domestic enplaned passengers and RPKs, respectively. The final linear GAPAXDE<sub>lin</sub> and GARPXSDE<sub>lin</sub> models comprise ten inputs.

$$GAPAXDE_{lin} = -36,620,320.54 + 5,070.49X_1 + 2.22X_2 - 15.38X_3 + 1,679.57X_4 + 286,641.85X_5 + 254,054.87X_6 + 4.75X_7 - 286,163.96X_8 - 312,420.85X_9 + 176,322.4X_{10} \quad (13)$$

$$GARPXSDE_{lin} = -35,915,376.4 + 3,790.49X_1 + 1.82X_2 + 18.35X_3 + 3,453.36X_4 + 343,252.11X_5 + 332,361.78X_6 - 0.85X_7 - 446,404.39X_8 - 384,392.98X_9 + 232,552.8X_{10} \quad (14)$$

where:

$X_1$  is Australia's real best discount economy airfare;  $X_2$  means Australia's population size;  $X_3$  is Australia's real GDP;  $X_4$  refers to Australia's unemployment size;  $X_5$  is Australia's real interest rates;  $X_6$  means world jet fuel prices;  $X_7$  recorded bed capacities at Australia's tourist accommodation establishments;  $X_8$  is Ansett's collapse (Dummy 1);  $X_9$  refers to Sydney Olympic Games (Dummy 2); and  $X_{10}$  are Commonwealth Games held in Melbourne (Dummy 3).

The GA selected the optimum variables for both the GAPAXDE<sub>quad</sub> and GARPXSDE<sub>quad</sub> quadratic models, and these are presented in the following equations:



$$\begin{aligned}
 GAPAXDE_{quad} = & -3,467,294.39 + 3,360,437.25X_6 - \\
 & 0.28X_1X_3 + 29.96X_1X_4 + 0.00097X_2X_4 - 9.13X_4^2 + \\
 & 3,548.08X_1X_5 + 6.44X_3X_5 - 1,243.83X_4X_5 - \\
 & 32,450.09X_5^2 - 1.08X_2X_6 + 29.58X_3X_6 - 545,127.83X_5X_6 + \\
 & 0.057X_1X_7 + 21.06X_6X_7 + 0.39X_3X_8
 \end{aligned}
 \tag{15}$$

$$\begin{aligned}
 GARPXSDE_{quad} = & 34,563,862.38 - 410,296.21X_1 - \\
 & 4.14X_2 + 155.09X_3 - 2,014,127.33X_6 + 0.04X_1X_2 - \\
 & 1.39X_1X_3 - 0.0000081X_3^2 + 0.05X_3X_4 - 920.51X_4X_5 - \\
 & 6,457.99X_4X_6 - 591,583.56X_5X_6 + 2.73X_5X_7 + 14.78X_6X_7
 \end{aligned}
 \tag{16}$$

Following the training procedure, which produced the weighting factors of the GA models, the testing procedure was performed using the 13 out-of-sample datasets to verify and test the accuracy, reliability, and the robustness of the GA models. The relative errors between the observed and estimated data for the two forms of GAPAXDE and GARPXSDE models (both linear and quadratic function forms) are presented in Tables 1 and 2, respectively. Table 1 compares the relative error between the actual and forecasted values in the testing phase of the GAPAXDE linear and quadratic models. The obtained average relative error for the GAPAXDE linear and quadratic models is 4.89 and 2.55%, respectively.

Similar to Table 1, Table 2 compares the relative error between actual and forecasted values in the testing phase of the GARPXSDE linear and quadratic models. The obtained

average relative error for the GARPXSDE linear and quadratic models is 4.89 and 2.23%, respectively.

Table 3 presents MAE, MAPE, MSE, and RMSE in training as well as out-of-sample testing dataset of the GAPAXDE linear and quadratic models for forecasting Australia’s domestic airline enplaned passengers. These results show that the GAPAXDE quadratic models performed better than the linear models during both training and testing phase as measured by MAE, MAPE, MSE, and RMSE. In the testing phase, where out-of-sample dataset was used to forecast Australia’s domestic airline enplaned passenger demand, the MAPE value of the GAPAXDE linear and quadratic models was 4.89 and 2.55%, respectively.

Table 4 presents MAE, MAPE, MSE, and RMSE in training as well as out-of-sample testing dataset of the GARPXSDE linear and quadratic models for forecasting Australia’s domestic airline RPKs. The GARPXSDE quadratic models performed better than the linear models during both training and testing phases. In the testing phase, the MAPE value of the GARPXSDE linear and quadratic models was 4.89 and 2.23%, respectively.

This study further employed a hypothesis test to give an indication if the difference between the models utilised was in fact statistically significant. Since the same 13 out-of-sample datasets were used for forecasting in all models, the paired t-test assuming unequal variance was used to assess

**Table 1.** A comparison of the GAPAXDE linear and quadratic modelling results with the observed data for the testing period.

Testing data	Actual PAX	Linear model	Relative error (%)	Quadratic model	Relative error (%)
1	13,532,828	13,757,250	1.66	13,467,576	0.48
2	13,535,413	14,016,870	3.56	13,291,868	1.80
3	13,459,758	14,113,293	4.86	13,215,146	1.82
4	13,561,072	14,216,017	4.83	14,028,957	3.45
5	13,649,165	14,206,416	4.08	13,699,900	0.37
6	13,799,188	14,375,221	4.17	13,872,894	0.53
7	14,013,059	14,503,786	3.50	13,476,807	3.83
8	14,072,782	14,710,848	4.53	13,972,716	0.71
9	14,154,745	14,789,210	4.48	14,110,082	0.32
10	14,227,970	15,075,341	5.96	14,644,886	2.93
11	14,282,647	15,214,615	6.53	14,722,852	3.08
12	14,343,780	15,399,232	7.36	15,347,930	7.00
13	14,355,000	15,516,248	8.09	15,331,447	6.80
MAPE (%)			4.89	2.55	

the forecasting accuracy of the respective models and also to test the hypothesis ( $H_0$ ) that there is not a significant difference in the forecasting accuracy of the GA linear and quadratic models (Razi and Athappilly 2005; Zaefizadeh *et al.* 2011). That is:

- $H_0: \mu_{linear} \leq \mu_{quadratic}$
- $H_1: \mu_{linear} > \mu_{quadratic}$

The results of t-tests are presented in Table 5, which shows that the p-values for both the GAPAXDE and GARPKSDE tests are less

than 0.05, therefore,  $H_{01}$  and  $H_{02}$  are rejected. This implies that the average forecasting error of the GA linear models is statistically significantly different from the average forecasting error of GA quadratic models (both GAPAXDE and GARPKSDE) at the 95% confidence interval of the difference. The results also indicated that the forecasting error of GA linear models is higher than that of the GA quadratic models. These results also confirm that the GA quadratic forms are superior to the linear forms when used to forecast Australia's domestic PAX and RPKs, respectively.

**Table 2.** A comparison of the GARPKSDE linear and quadratic modelling results with the observed data for the testing period.

Testing data	Actual RPKs	Linear model	Relative error (%)	Quadratic model	Relative error (%)
1	15,437,103	15,737,744	1.95	15,288,311	0.96
2	15,554,857	16,187,217	4.07	14,958,997	3.83
3	15,580,866	16,327,747	4.79	15,260,316	2.06
4	15,788,103	16,508,532	4.56	16,030,152	1.53
5	15,942,594	16,550,531	3.81	15,980,455	0.24
6	16,129,413	16,838,573	4.40	15,995,064	0.83
7	16,389,404	16,915,031	3.21	16,271,617	0.72
8	16,469,044	17,174,398	4.28	16,446,808	0.14
9	16,579,654	17,317,906	4.45	16,544,018	0.21
10	16,688,541	17,619,437	5.58	17,338,501	3.89
11	16,767,381	17,852,751	6.47	17,438,574	4.00
12	16,866,762	18,148,297	7.60	17,885,524	6.04
13	16,913,071	18,336,509	8.42	17,676,766	4.52
MAPE (%)			4.89	2.23	

**Table 3.** Performance index of GAPAXDE linear and quadratic models for training and testing (out of sample) data set.

Performance index	GAPAXDE linear model		GAPAXDE quadratic model	
	Training data	Testing data	Training data	Testing data
MAE	230,821	685,149	138,016	358,803
MAPE	2.87%	4.89%	1.66%	2.55%
MSE	$98.3 \times 10^9$	$529 \times 10^9$	$32.4 \times 10^9$	$229 \times 10^9$
RMSE	313,513	727,490	179,957	478,722

**Table 4.** Performance index of GARPKSDE linear and quadratic models for training and testing (out of sample) data set.

Performance index	GARPKSDE linear model		GARPKSDE quadratic model	
	Training data	Testing data	Training data	Testing data
MAE	254,561	800,606	152,326	366,056
MAPE	2.98%	4.89%	1.78%	2.23%
MSE	$117 \times 10^9$	$729 \times 10^9$	$38.4 \times 10^9$	$236 \times 10^9$
RMSE	341,975	853,569	195,883	485,775

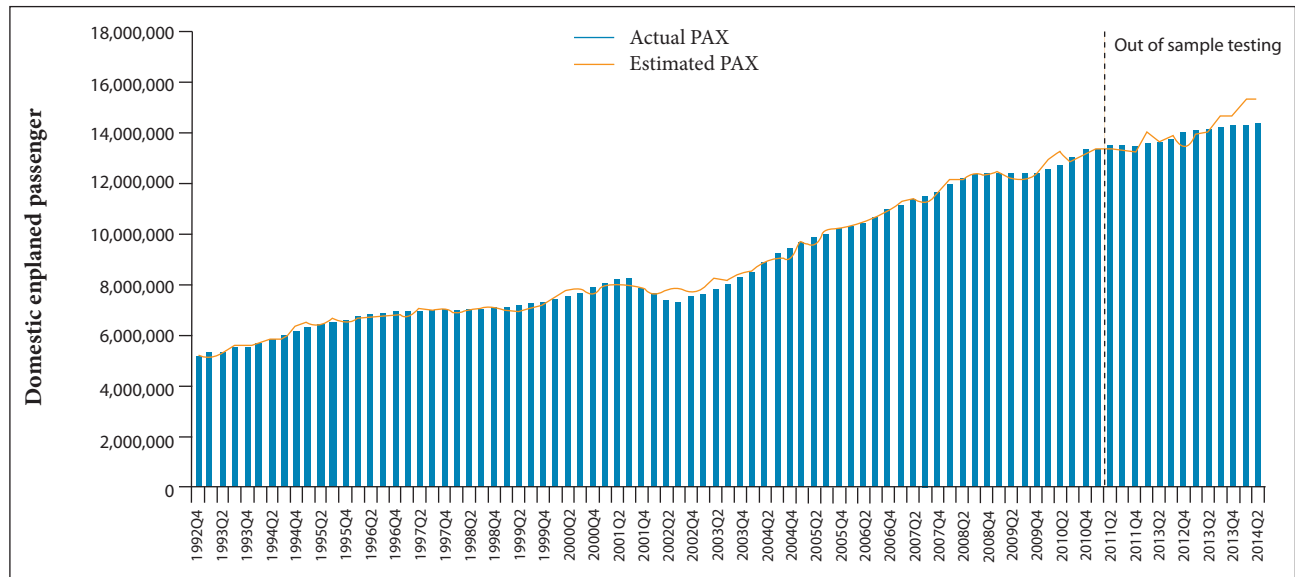
Australia's actual quarterly domestic airline and forecast enplaned passengers, during the period from Quarter 4 1992 to Quarter 2 2014, are plotted and shown in Fig. 2. Finally,

Australia's actual quarterly domestic airline and forecast RPKs, during the period from Quarter 4 1992 to Quarter 2 2014, are plotted and shown in Fig. 3.

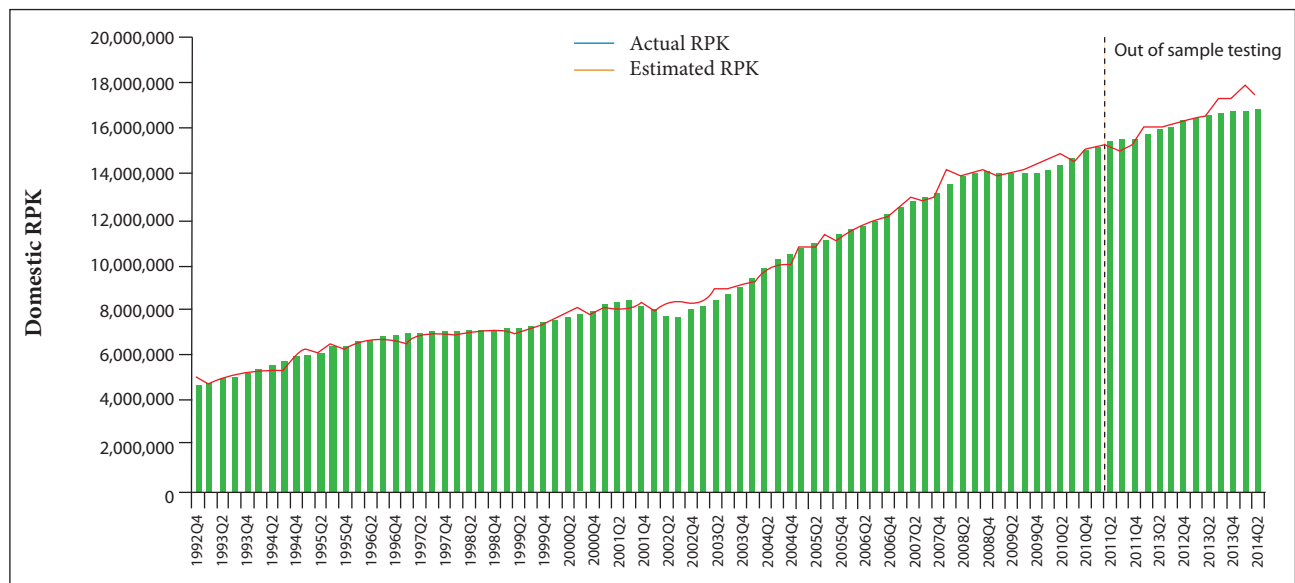
**Table 5.** Results of paired t-test.

Test	t-stat	p-value	Conclusion
GAPAXDE $H_{01}$ : linear versus quadratic	5.70	< 0.01	$\mu_{linear} > \mu_{quadratic}$
GARPKSDE $H_{02}$ : linear versus quadratic	7.55	< 0.01	$\mu_{linear} > \mu_{quadratic}$

$\mu_{linear} > \mu_{quadratic}$  are mean forecasting errors of GA linear and quadratic models, respectively.



**Figure 2.** A comparison of Australia's actual and forecast quarterly domestic airline enplaned passengers (GAPAXDE model).



**Figure 3.** A comparison of Australia's actual and forecast quarterly domestic airline RPKs (GARPKSDE model).

## CONCLUSIONS

This study has developed and empirically examined GA models for forecasting Australia's quarterly domestic airline passenger demand (GAPAXDE and GARPXSDE models), as measured by enplaned passengers and RPKs. Two mathematical forms — linear and quadratic — were developed and tested in the study.

The dataset used in the study covers the period from Quarter 4 1992 to Quarter 2 2014 as this represented the period of time in which data for each of the variables selected for examination in the modelling was available. Data was divided into two sets, training and testing dataset; 74 training datasets were used to estimate the weighting factors of the GA model and 13 out-of-sample datasets were used for testing the robustness of the GA models. The GA parameters used in this study comprised population size ( $n$ ): 200; the generation number: 1,000; and mutation rate: 0.01. The modelling results have shown that both quadratic GAPAXDE and GARPXSDE models are more accurate, reliable and have greater predictive capability as compared to the linear form. The out-of-sample testing dataset MAPE

of GAPAXDE and GARPXSDE quadratic models is 2.55 and 2.23%, respectively.

As air travel markets often compromise regional and international market segments, a suggestion for future research would be to empirically examine the models to estimate a country's regional or international passenger air travel demand. In the event that the data is available, the models could also be used to forecast LCC and FSNC air travel demand.

Future research will investigate the implementation of other meta-heuristics to facilitate a comparison such that the optimum method can be identified for use in the aviation industry. This work will also be compared with auto-regression methods, and then hybrid meta-heuristics and auto-regression methods. The use of auto-regression facilitates the lagging of the independent variables to investigate quasi time-related effects. This then captures delays by a quarter or more on the dependent of the independent variables. The combination with meta-heuristics to search for which variables to include and the possibility to also search for lagged variables may greatly improve the forecasting power of the model.

## REFERENCES

- Abed SY, Ba-Fail AO, Jasimuddin SM (2001) An econometric analysis of international air travel demand in Saudi Arabia. *J Air Transp Manag* 7(3):143-148. doi: 10.1016/S0969-6997(00)00043-0
- Aderamo AJ (2010) Demand for air transport in Nigeria. *J Econ* 1(1):23-31.
- Akgüngör AP, Doğan E (2009) An artificial intelligent approach to traffic accident estimation: model development and application. *Transport* 24(2):135-142. doi: 10.3846/1648-4142.2009.24.135-142
- Alekseev KPG, Seixas JM (2009) A multivariate neural forecasting modeling for air transport - preprocessed by decomposition: a Brazilian application. *J Air Transp Manag* 15(5):212-216. doi: 10.1016/j.jairtraman.2008.08.008
- Amjadi MH, Nezamabadi-Pour H, Farsangi MM (2010) Estimation of electricity demand of Iran using two heuristic algorithms. *Energy Conserv Manage* 51(3):493-497. doi: 10.1016/j.enconman.2009.10.013
- Azadeh A, Saberi M, Anvari M, Azaron A, Mohammadi M (2011) An adaptive network based fuzzy inference system-genetic algorithm clustering ensemble algorithm for performance assessment and improvement of conventional power plants. *Expert Syst Appl* 38(3):2224-2234. doi: 10.1016/j.eswa.2010.08.010
- Azadeh A, Ghaderi SF, Tarveredian S, Saberi M (2007) Integration of artificial neural networks and genetic algorithm to predict electrical energy consumption. *Appl Math Comput* 186(2):1731-1741. doi: 10.1016/j.amc.2006.08.093
- Azadeh A, Saberi M, Ghorbani S (2010) An ANFIS algorithm for improved forecasting of oil consumption: a case study of USA, Russia, India and Brazil. *J Sci Ind Res* 69(3):194-203.
- Ba-Fail AO, Abed SY, Jasimuddin SM (2000) The determinants of domestic air travel demand in the kingdom of Saudi Arabia. *J Air Transport World Wide* 5(2):72-86.
- Baker D, Donnet T (2012) Regional and remote airports under stress in Australia. *Res Trans Bus Manage* 4:37-43. doi: 10.1016/j.rtbm.2012.06.011
- Belobaba PP (2009) Overview of airline economics, markets and demand. In: Belobaba P, Odoni A, Barnhart C, editors. *The global airline industry*. Chichester: John Wiley & Sons. p. 47-72.
- Bhadra D (2003) Demand for air travel in the United States: bottom-up econometric estimation and implications for forecasts by origin and destination pairs. *J Air Transp* 8(2):19-56.
- Bieger T, Wittmer A (2006) Air transport and tourism - perspectives and challenges for destinations, airlines and governments. *J Air Transp Manag* 12(1):40-46. doi: 10.1016/j.jairtraman.2005.09.007
- Blinova TO (2007) Analysis of Possibility of Using Neural Network to Forecast Passenger Traffic Flows in Russia. *Aviation* 11(1):28-34. doi: 10.1080/16487788.2007.9635952

- Bureau of Transport and Communication Economics (1994) International aviation: trends and issues. Report 86. Canberra: Australian Government Publishing Service.
- Carvajal-Rodríguez A, Carvajal-Rodríguez F (2009) Population extinction in genetic algorithms: application in evolutionary studies. In: Omatu S, Rocha MP, Bravo J, Fernández F, Corchado E, Bustillo A, Corchado JM, editors. Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living. Lecture notes in Computer Science. Vol. 5518. Berlin: Springer. p. 1028-1034.
- Cervantes L, Costillo O (2013) Genetic optimization of membership functions in modular fuzzy controllers for complex problems. In: Castillo O, Melin P, Kacprzyk J, editors. Recent advances on hybrid intelligent systems. Berlin: Springer-Verlag. p. 51-62.
- Chen MS, Ying LC, Pan MC (2010) Forecasting tourist arrivals by using the adaptive network-based fuzzy inference system. *Expert Syst Appl* 37(2):1185-1191. doi: 10.1016/j.eswa.2009.06.032
- Clark DE, McGibany JM, Myers A (2009) The effects of 9/11 on the airline travel industry. In: Morgan MJ, editor. The impact of 9/11 on business and economics: the business of terror - the day that changed everything? London: Palgrave Macmillan. p. 76-86.
- Coelho RF, Méndez Echenagucia T, Pugnale A, Richardson JN (2014) Genetic algorithms for structural design. In: Adriaenssens S, Block P, Veenendaal D, Williams C, editors. Shell structures for architecture: form finding and optimization. Abingdon: Routledge. p. 290-294.
- Cook A (2007) European air traffic management: principles, practice, and research. Aldershot: Ashgate Publishing.
- Dempsey PS, Gesell LE (1997) Airline management: strategies for the 21st century. Chandler: Coast Aire Publications.
- Doganis R (2009) Flying off course: airline economics and marketing. 4th edition. Abingdon: Routledge.
- Easdown G, Wilms P (2002) Ansett: the collapse. Melbourne: Lothian Books.
- Forsyth P (2003) Low-cost carriers in Australia: experiences and impacts. *J Air Transp Manag* 9(5):277-284. doi: 10.1016/S0969-6997(03)00035-8
- Gen M, Cheng R (1997) Genetic algorithms and engineering design. New York: John Wiley & Sons.
- Gesell LE (1993) The administration of public airports. 3rd edition. Chandler: Coast Aire Publications.
- Ghanbari A, Kzemi SMR, Mehmanpazir F, Nakhostin MM (2013) A cooperative ant colony optimization-genetic algorithm approach for construction of energy demand forecasting knowledge-based expert systems. *Knowl-Based Syst* 39:194-206. doi: 10.1016/j.knosys.2012.10.017
- Godinho P, Silva M (2014) Genetic, memetic and electromagnetism-like algorithms: applications in marketing. In: Moutinho L, Bigné E, Manrai AK, editors. The Routledge Companion to the future of marketing. Abingdon: Routledge. p. 393-416.
- Goldberg DE (1989) Genetic algorithms in search, optimization, and machine learning. Reading: Addison-Wesley.
- Haldenbilen S, Ceylan H (2005) Genetic algorithm approach to estimate transport energy demand in Turkey. *Energy Policy* 33(1): 89-98. doi: 10.1016/S0301-4215(03)00202-7
- Hanlon P (2007) Global airlines: competition in a transnational industry. 3rd edition. Oxford: Butterworth-Heinemann.
- Hensher DA, Brewer AM (2002) Going for gold at the Sydney Olympics: how did transport perform?. *Transport Rev* 22(4):381-399. doi: 10.1080/01441640110121112
- Hernández-López M, Cáceres-Hernández JJ (2007) Forecasting tourists' characteristics by a genetic algorithm with a transition matrix. *Tourism Manage* 28(1):290-297. doi: 10.1016/j.tourman.2005.11.016
- Holland JH (1975) Adaption in natural and artificial systems. Ann Arbor: University of Michigan Press.
- Holloway S (2008) Straight and level: practical airline economics. 3rd edition. Aldershot: Ashgate Publishing.
- Hong WC, Dong Y, Chen LY, Wei SY (2011) SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. *Appl Soft Comput* 11(2):1881-1890. doi: 10.1016/j.asoc.2010.06.003
- Hooper K, van Zyl M (2011) Australia's tourism industry. *Bulletin, December Quarter 2011:23-32*; [accessed 2015 Nov 11]. <http://www.rba.gov.au/publications/bulletin/2011/dec/pdf/bu-1211-3.pdf>
- Hu C (2002) Advanced tourism demand forecasting: artificial neural network and Box-Jenkins modelling (PhD thesis). West Lafayette: Purdue University.
- Hurley S, Moutinho L, Witt S (1998) Genetic algorithms for tourism marketing. *Ann Tourism Res* 25(2):498-514. doi: 10.1016/S0160-7383(97)00099-6
- International Air Transport Association (2008) Air travel demand. IATA Economics Briefing No. 9; [accessed 2015 Jan 20]. [http://www.iata.org/whatwedo/Documents/economics/air\\_travel\\_demand.pdf](http://www.iata.org/whatwedo/Documents/economics/air_travel_demand.pdf)
- International Air Transport Association (2015) Airline profitability improves with falling oil prices; [accessed 2015 Aug 30]. <http://www.iata.org/pressroom/pr/Pages/2014-12-10-01.aspx>
- International Civil Aviation Organization (2006) Manual on air traffic forecasting. Document No. 8991 AT722/3. 3rd edition. Montreal: ICAO.
- Jirong G, Mingchang Z, Liuguangyan J (2011) Housing price forecasting based on genetic algorithm and support vector machine. *Expert Syst Appl* 38(4):3383-3386. doi: 10.1016/j.eswa.2010.08.123
- Jones KO, Romil S (2004) Selection of an optimum feed profile using genetic algorithms. Proceedings of the 9th IFAC International Symposium; Nancy, France.
- Koo TTR, Tan DT, Duval DT (2013) Direct air transport and demand interaction: a vector error-correction model approach. *J Air Transp Manag* 28:14-19. doi: 10.1016/j.jairtraman.2012.12.005
- Kopsch F (2012) A demand model for domestic air travel in Sweden. *J Air Transp Manag* 20:46-48. doi: 10.1016/j.jairtraman.2011.11.006
- KPMG (2006) Economic Impact Study of the Melbourne 2006 Commonwealth Games Post-event Analysis. Office of Commonwealth Games Coordination; [accessed 2014 Apr 15]. [http://www1.dvc.vic.gov.au/ocgc/News/econ\\_impact\\_report.pdf](http://www1.dvc.vic.gov.au/ocgc/News/econ_impact_report.pdf)
- Kunt MM, 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. doi: 10.3846/16484142.2011.635465

- Li G (2010) The nature of leisure travel demand. In: Graham A, Papatheodorou A, Forsyth P, editors. *Aviation and tourism: implications for leisure travel*. Aldershot: Ashgate Publishing. p. 7-20.
- Madden JR (2002) The economic consequences of the Sydney Olympics: the CREA/Arthur Andersen Study. *Curr Issues Tourism* 5(1):7-21. doi: 10.1080/13683500208667904
- McKnight P (2010) Airline economics. In: Wald A, Fay C, Gleich R, editors. *Introduction to aviation management*. Münster: LIT Verlag. p.25-53.
- Nolan JL, editor (1996) *Australia business: the portable encyclopaedia for doing business with Australia*. San Rafael, CA: World Trade Press.
- Organization for Economic Cooperation and Development (2014) *OECD Tourism Trends and Policies 2014*. Paris: OECD Publishing.
- Ozturk HK, Ceylan H, Canyur OE, Hepbasli A (2005) Electricity estimation using genetic algorithm approach: a case study of Turkey. *Energy* 30(7):1003-1012. doi: 10.1016/j.energy.2004.08.008
- Prideaux B (2003) The need to use disaster planning frameworks to respond to major tourism disasters: analysis of Australia's response to major tourism disasters in 2001. *J Trav Tourism Market* 15(4):281-298. doi: 10.1300/J073v15n04\_04
- Razi AM, Athappilly K (2005) A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Syst Appl* 29(1):65-74. doi: 10.1016/j.eswa.2005.01.006
- Rotshtein AP, Rakytyanska HB (2012) *Fuzzy evidence in identification, forecasting and diagnosis*. Heidelberg: Springer.
- Sineglazov V, Chumachenko E, Gorbatyuk V (2013) An algorithm for solving the problem of forecasting. *Aviation* 17(1):9-13. doi: 10.3846/16487788.2013.777219
- Sivrikaya O, Tunç E (2013) Demand forecasting for domestic air transportation in Turkey. *Open Transport J* 7:20-26.
- Srisaeng P (2015) *Utilizing advanced modelling approaches for forecasting air travel demand: a case study of Australia's domestic low cost carriers* (PhD thesis). Melbourne: RMIT University.
- Srisaeng P, Baxter GS, Wild G (2014) The evolution of low cost carriers in Australia. *Aviation* 18(4):203-216. doi: 10.3846/16487788.2014.987485
- 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. *Construct Build Mater* 62:102-108. doi: 10.1016/j.conbuildmat.2014.03.041
- Tsekeris T (2009) Dynamic Analysis of Air Travel Demand in Competitive Island Markets. *J Air Transp Manag* 15(6):267-273. doi: 10.1016/j.jairtraman.2008.11.008
- Vasigh B, Tacker T, Fleming K (2008) *Introduction to air transport economics: from theory to applications*. Aldershot: Ashgate Publishing.
- Wensveen JG (2011) *Air transportation: a management perspective*. 7th edition. Farnham: Ashgate Publishing.
- Yetilmezsoy K, Fingas M, Fieldhouse B (2011) An adaptive neuro-fuzzy approach for modeling of water-in-oil emulsion formation. *Colloid Surface A* 389(1-3):50-62. doi: 10.1016/j.colsurfa.2011.08.051
- Young SB, Wells AT (2011) *Airport planning and management*. 6th edition. New York: McGraw-Hill.
- Zaefizadeh M, Khayatnezhad M, Gholamin R (2011) Comparison of multiple linear regressions (MLR) and artificial neural network (ANN) in predicting the yield using its components in the hullless barley. *Am Eurasian J Agric Environ Sci* 10(1):60-64.