

FORECASTING AUSTRALIA'S DOMESTIC LOW COST CARRIER PASSENGER DEMAND USING A GENETIC ALGORITHM APPROACH

Panarat SRISAENG¹, Steven RICHARDSON², Glenn BAXTER³, Graham WILD⁴

^{1, 3, 4} School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University, Melbourne, Australia 3001

²School of Engineering, Edith Cowan University, Joondalup, Western Australia, Australia 6027
E-mail: s3125221@student.rmit.edu.au¹ (corresponding author)

Received 23 October 2014; accepted 24 March 2016



Panarat SRISAENG

Education: Bachelor of Economics, Chulalongkorn University, Bangkok, Thailand, 1993. Master of Business Economics, Kasetsart University, Bangkok, Thailand, 1998. Affiliations and functions: PhD (Candidate) Aviation, RMIT University, School of Aerospace, Mechanical and Manufacturing Engineering. Research interests: low cost airlines management, demand model for air transportation, demand forecasting for air transportation.



Steven RICHARDSON, PhD.

Education: Bachelor of Science, The University of Western Australia, Perth, Western Australia, Australia, 2008. PhD, The University of Western Australia, Perth, Western Australia, Australia, 2008. Affiliations and functions: Senior Lecturer, School of Science, Edith Cowan University. Research interests: numerical methods/analysis, optimisation simulation modelling (applied to iron ore mining), simulation of tyre usage in iron ore mining, simulation of repair shop in iron ore mining.



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 Manager of Undergraduate Aviation Programs, RMIT University, School of Engineering. Research interests: air cargo handling and operations, airport operations and sustainability, supply chain management.



Graham WILD, PhD.

Education: 2001–2004, Bachelor of Science (Physics and Mathematics), Edith Cowan University. 2004–2005, Bachelor of Science Honours (Physics), Edith Cowan University. 2008, Graduate Certificate (Research Commercialisation), Queensland University of Technology. 2006–2008, Master of Science and Technology (Photonics and Optoelectronics), the University of New South Wales. 2006–2010, PhD (Engineering), Edith Cowan University. Affiliations and functions: 2013, Aviation Program Manager, RMIT University, School of Engineering. 2010, Postdoctoral research

associate, Photonics Research Laboratory, Edith Cowan University. 2011–2012, Lecturer of Aviation Systems, Edith Cowan University. 2012, Senior Lecturer in Aerospace and Aviation, RMIT University, School of Engineering.

Research interests: aircraft systems, aerospace vehicle structural health monitoring, optical fibres, sensing and instrumentation, sustainable aviation, aviation technology. aircraft control with an emphasis on interactions between human and machine.

Abstract. This study has proposed and empirically tested for the first time Genetic Algorithm (GA) models for forecasting Australia's domestic low cost carriers' demand, as measured by enplaned passengers (GAPAXDE Model) and revenue passenger kilometres performed (GARPKSDE Model). Data was divided into training and testing data sets, 36 training data sets were used to estimate the weighting factors of the GA models and 6 data sets were used for testing the robustness of the GA models. The genetic algorithm parameters used in this study comprised population size (n): 1000, the generation number: 200, and mutation rate: 0.01. The modelling results have shown that both the linear GAPAXDE and GARPKSDE models are more accurate, reliable, and have a slightly greater predictive capability compared to the quadratic models. The overall mean absolute percentage error (MAPE) of the GAPAXDE and GARPKSDE models are 3.33 per cent and 4.48 per cent, respectively.

Keywords: Australia, forecasting method, genetic algorithm (GA); low cost carriers; air transport.

1. Introduction

Australia's airline industry was born on connecting regional communities to the country's major cities (Baker, 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 31 per cent market share, with the two major incumbent LCCs being Jetstar and Tiger Airways.

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 assisting 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).

Despite the significance of Australia's low cost carrier domestic airline market sector, there has been no previously reported study that has developed and empirically examined genetic algorithm-based models for forecasting Australia's domestic low cost carrier passenger demand. The primary objective of this study is therefore to develop new genetic algorithm-based models to forecast Australia's LCCs passenger demand and also to

identify whether the GA approach is a useful tool for this application. Therefore, various forms of the mathematical expressions were proposed and tested. Genetic algorithm enplaned passengers (GAPAXDE) and genetic algorithm revenue passenger kilometres performed¹ (GARPKSDE) are proposed to forecast Australia's LCC quarterly enplaned passengers and revenue passenger kilometres performed, respectively.

2. Traditional air travel demand forecasting approaches

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, 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). In addition, forecasts assist airlines in their decision-making regarding the development of infrastructure facilities, thereby enhancing services provided to passengers (Abed *et al.* 2001).

In the air transport industry, many service providers and government regulatory agencies follow the International Civil Aviation Organization's (ICAO) *Manual on Air Traffic Forecasting*. This manual was originally developed in 1985 using traditional modelling techniques

¹ Airline passenger traffic can be measured in two ways; the number of passengers carried and also by revenue passenger kilometres performed (RPKs) (Belobaba 2009; Holloway 2008).

(Alekseev, Seixas 2009). Indeed, in the past, multiple linear regression (MLR) models have been generally used to forecast air traffic demand (see, for example, Abed *et al.* 2001; Aderamo 2010; Ba-Fail *et al.* 2000; Bhadra 2003; International Civil Aviation Organization 2006; Kopsch 2012; Sivrikaya, 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: 502). 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, 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 A., Carvajal-Rodríguez F. 2009).

This 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, Cáceres-Hernández 2007; Hong *et al.* 2011), tourism marketing (Hurley *et al.* 1998), traffic accident severity prediction (Akgüngör, Doğan 2009; Kunt *et al.* 2011), and transport energy demand prediction (Haldenbilen, Ceylan 2005). In addition, Sineglazov *et al.* (2013) have proposed a GA forecasting method for solving the problems of forecasting experienced in the aviation industry. These authors have also noted that their GA may be applicable for forecasting regional aviation facilities and other industrial sectors that have demand patterns similar to those experienced by airlines.

3. 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, Doğan 2009). GAs are therefore similar to the natural evolution process where a 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, Romil 2004).

The principal strength of GAs is their adaptive and self-organizing capabilities. These abilities enable GAs to quickly solve difficult problems through three evolutionary mechanisms: (1) selection, (2) crossover, and (3) mutation (Hu 2002).

The basic operations of GAs include selection, a crossover of genetic information between reproducing parents and a mutation of genetic information which affect 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, Doğan 2009).

3.1 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, Cheng 1997). GAs generates 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, Silva 2014: 395; Hurley *et al.* 1998). 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* (Gen, Cheng 1997). 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 (1) merging two chromosomes from a current generation using a crossover operator, or (2) by modifying a chromosome using a mutation operator (Gen, Cheng 1997: 2). A new generation is formed by (1) selecting, according to fitness values, some of the parents and the off-spring whilst (2) rejecting others so 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, Cheng 1997: 2).

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.

The *fitness function*, (that is, minimum sum of squared errors [SSE]) $F(x)$, is presented as follows:

$$\min_s F_i(x) = \sum_j^m s_j (t_i - td_i)^2, \quad (1)$$

where t_i and td_i are the actual and estimated value, respectively, m is the number of observations, and $s = \{S_j\}$ is the vector of weighting factors. The GA process is illustrated in Figure 1.

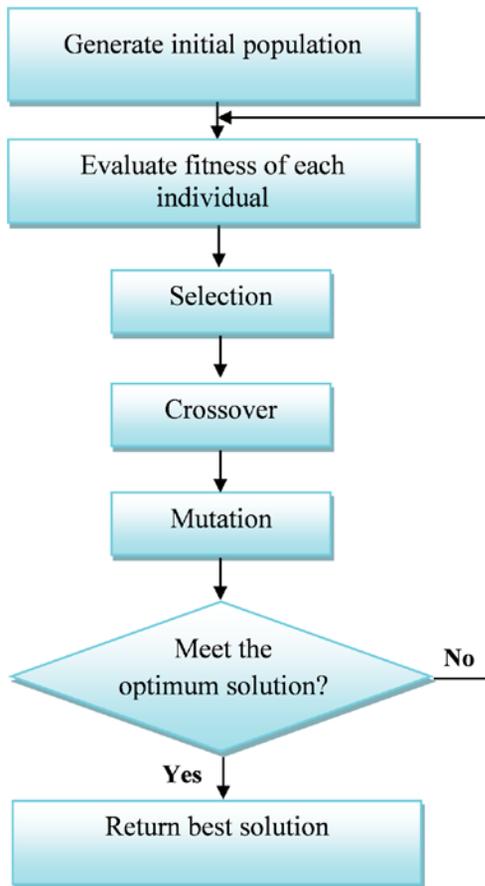


Fig. 1. The genetic algorithm process
Source: adapted from (Amjadi *et al.* 2010).

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).

When using a GA, it is a requirement to select chromosomes from the current population for reproduction. The selection procedure picks out 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: 502).

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 (Hurley *et al.* 1998). 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 chromosome set (Akgüngör, 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, 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: 1006).

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, 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 of a switching of a 0–1 or vice versa (Ozturk *et al.* 2005: 1006). 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, Doğan 2009):

Offspring	10101110 1101010
Mutated Offspring	10101110 0101010

4. The genetic algorithm models for forecasting Australia's domestic LCCs passenger demand

4.1 The GAPAXDE and GARPKSDE data and variables selection

In this study, genetic algorithm prediction models (GARPKSDE) for Australia's domestic low cost carrier (LCC) enplaned passenger (GAPAXDE) and Australia's LCC domestic revenue passenger kilometres performed (RPKs) have been proposed. During the model development process, eleven variables were selected as the independent variables in the two GA models: Australia's real GDP, Australia's real GDP per capita, Australia's real best discount air fares, Australia's population size, Australia's unemployment size, Australia's tourist accommodation establishments recorded bed capacities, world jet fuel prices, and Australia's real interest rates. Four dummy variables were also included in the GA models. The first dummy variable signified the impact of the evolving Virgin Australia business model from a low cost carrier (LCC) 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 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 (DUMMY 2) 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) from the middle of 2007 to March 2009. The GFC had a negative impact on both Australia's financial markets and the Australian economy. Importantly, it had a direct effect on Australian households' income (Debelle 2009), and hence, their ability and willingness to travel.

The fourth dummy variable accounted for the impact of the Commonwealth Games held in Melbourne from 15 to March 26, 2006. The 2006 Melbourne Commonwealth Games was the largest sporting and community event held in Victoria's history. The Commonwealth Games provided substantial economic benefits for the State of Victoria, and for the tourism and airline industry (KPMG 2006).

The availability of a consistent data set allows the use of quarterly data for the period 2002 to 2012. The data used in the GA 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 establishments² 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 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).

4.2. The GAPAXDE and GARPKSDE genetic algorithm process

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 will be considered:

$$\hat{y} = \sum_{i=1}^n w_i a_i X_i ; \quad (2)$$

$$\hat{y} = \sum_{\substack{i=1 \\ i \in I}}^n w_i a_i \ln(X_i) + \sum_{\substack{i=1 \\ i \notin I}}^n w_i a_i X_i ; \quad (3)$$

$$\hat{y} = \sum_{i=1}^n w_i a_i X_i + \sum_{i=1}^n \sum_{j=1}^i w_{ij} a_{ij} X_i X_j ; \quad (4)$$

where, I is the set of variables which only take positive values (that is, not zero or negative).

The coefficient component w_i indicates whether the variable X_i is included in the model, where $w_i = 1$ if X_i is

² Based on Australian tourist accommodation establishments with 15 rooms or more.

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 $W_k = \{\mathbf{w} : \mathbf{w} \cdot \mathbf{1} = k\}$.

Once \mathbf{w} is specified, the values of $\{a_i : i = 1, 2, \dots, n\}$ (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 $\{y_i : i = 1, 2, \dots, m\}$, and the corresponding predicted values $\{\hat{y}_i : i = 1, 2, \dots, m\}$ (i.e. least squares). That is, denoting the vector of model coefficients as \mathbf{a} , we minimise:

$$LS(\mathbf{a} | \mathbf{w}) = \sum_{i=1}^m (y_i - \hat{y}_i)^2. \quad (5)$$

Objective function

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

$$\mathbf{w}^* = \arg \min_{\mathbf{w} \in W_K} \left[\min_{\mathbf{a}} LS(\mathbf{a} | \mathbf{w}) \right] \quad (6)$$

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, and $\binom{n^2 + 3n}{2K}$ for the quadratic model. This may make it prohibitive to exhaustively evaluate all models with K independent variables. One possible approach to determining a close to optimal set of independent variables is to utilise a meta-heuristic algorithm such as a GA.

The stages of the GA 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 $\mathbf{w} \in P_0$ of the initial population we define the measure of fitness $F(\mathbf{w})$ to be:

$$F(\mathbf{w}) = \min_{\mathbf{a}} LS(\mathbf{a} | \mathbf{w}). \quad (7)$$

2. Breed New Population Members for the 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} 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 $\mathbf{w} \in P_{i-1}$ for breeding each new solution is given by:

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

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

1. 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 onto the child solution).
2. 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 $\mathbf{w}_1(i) = 0, \mathbf{w}_2(i) = 1$.
3. 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. The 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 $\mathbf{w} \in P_{i-1}$ for elimination is given by:

$$\Pr(\mathbf{w}) = \frac{F(\mathbf{w})}{\sum_{\mathbf{w}_j \in P_{i-1}} F(\mathbf{w}_j)}. \quad (9)$$

5. Form a 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 pre-defined number of cycles, or until a pre-specified number of generations pass without improvement.

4.3. The 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. 2011). In this study, the first group of 36 data was used as the training set (85 per cent of the overall data), and the remaining 6 data group was 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): 1000;
- iterations (the generation number): 200;
- mutation rate: 0.01.

After applying the procedures of the GAPAXDE and GARPXSDE models, the following equations for forecasting Australia’s domestic LCCs enplaned passengers (PAX) and revenue paenger kilometres (RPKs) are obtained based on the minimum sum of squares error between the observed and estimated data.

The results of the GAPAXDE and GARPXSDE are given in the following equations:

$$GAPAXDE_{lin} = -291753.49 - 1204.25 \ln X_1 + 1484.72 \ln X_2 + 4854.94 \ln X_3 - 227.07 \ln X_5 + 1060.75 \ln X_6 + 120.71 \ln X_7 + 17075.24 \ln X_8 - 4326.67 X_9 + 13.29 X_{10} - 230.99 X_{11} - 157.26 X_{12}$$

$$GARPXSDE_{lin} = -386259.03 - 880.92 \ln X_1 + 5506.05 \ln X_2 + 4837.93 \ln X_3 - 625.67 \ln X_5 + 1118.91 \ln X_6 + 130.33 \ln X_7 + 21,272.44 \ln X_8 - 4398.22 X_9 - 145.25 X_{10} - 277.63 X_{11} - 211.04 X_{12}$$

The $GAPAXDE_{lin}$ and $GARPXSDE_{lin}$ present the linear logarithmic models for forecasting Australia’s domestic LCCs enplaned passengers and revenue passenger kilometres (RPKs), respectively. 11 inputs are selected for both $GAPAXDE_{lin}$ and $GARPXSDE_{lin}$ models: X_1 – Australia’s real best discount economy airfare, X_2 – Australia’s population size, X_3 – Australia’s real GDP, X_5 – Australia’s unemployment size, X_6 – world jet fuel prices, X_7 – Australia’s real interest rates, X_8 – recorded bed capacities at Australia’s tourist accommodation establishments, X_9 – dummy variable (Dummy 1) reflecting Virgin Australia changing business model, X_{10} – dummy variable (Dummy 2) accounting for the loss of capacity following the collapse of Ansett Australia, X_{11} – dummy variable (Dummy 3) reflecting the effects of the GFC and X_{12} – dummy variable (Dummy 4) reflecting the Commonwealth Games in 2006.

$$GAPAXDE_{quad} = -1065.28 - 79.38 X_1 - 0.005 X_3 - 4793.28 X_9 + 288538.5 X_{10} - 0.01 X_5^2 + 0.000001 X_2 X_3 - 0.004 X_3 X_6 + 3.7 X_5 X_6 + 0.0001 X_1 X_8 + 0.00001 X_5 X_8 + 14.8 X_2 X_{10}$$

$$GARPXSDE_{quad} = 18325.787 - 400.83 X_1 - 0.03 X_8 - 30.12 X_7^2 + 0.000004 X_2 X_3 - 0.04 X_1 X_5 + 1.86 X_1 X_6 - 0.003 X_2 X_7 + 0.0007 X_1 X_8 - 0.0000001 X_3 X_8 + 0.23 X_2 X_9 - 0.03 X_3 X_9$$

The optimum quadratic model for both $GAPAXDE_{quad}$ and $GARPXSDE_{quad}$ models incorporated 11 inputs in the model. This was similar to the GA linear models. Thus, the GA selected the optimum 11 variables for both the $GAPAXDE_{quad}$ and $GARPXSDE_{quad}$ models, as presented in the above equations.

Following the training procedure, which produced the weighting factors of the GA models, the testing procedure was performed using the remaining 6 data group 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, and linear and quadratic function forms are presented in Table 1 and Table 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 3.29 per cent and 5.01 per cent, respectively.

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	4,885.53	4,957.70	1.48	5,000.07	2.34
2	5,383.46	5,609.17	4.19	5,831.91	8.33
3	5,870.19	5,976.81	1.82	5,918.20	0.82
4	6,659.99	7,206.02	8.20	7,120.76	6.92
5	6,949.69	6,720.86	3.29	6,890.80	0.85
6	3,654.84	3,652.34	0.81	3,260.13	10.8
MAPE (%)			3.29	5.01	

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 3.45 per cent and 5.53 per cent, respectively.

Table 2. A comparison of the GARPXSDE 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	4,907.19	4,979.36	1.47	4,915.62	0.62
2	5,566.30	5,747.68	3.26	5,956.29	10.64
3	6,103.45	6,311.00	3.40	6,025.45	2.65
4	7,091.19	7,693.40	8.49	7,602.08	14.15
5	7,349.92	7,150.98	2.71	7,301.50	5.06
6	4,003.33	4,058.85	1.39	3,657.33	0.07
MAPE (%)			3.45	5.53	

Table 3 presents the mean average percentage error (MAPE) in training, testing and overall data set of both the GAPAXDE and GARPXSDE (linear and quadratic) models. During the training phase, both the GAPAXDE and GARPXSDE quadratic models performed better than the linear models as measured by both models mean absolute

percentage error (MAPE) values. The MAPE value of the GAPAXDE and GARPXSDE quadratic models was 2.38 per cent and 4.02 per cent, respectively. However, in the testing phase, both the GAPAXDE and GARPXSDE linear models performed better than the quadratic models. The MAPE value of the GAPAXDE and GARPXSDE linear models was 3.29 per cent and 3.45 per cent respectively, during the testing phase. The overall MAPE value of the linear GAPAXDE and GARPXSDE models is 3.33 per cent and 4.48 per cent as compared to 3.70 and 4.78 per cent for the GAPAXDE and GARPXSDE quadratic models. This suggests that the GA linear form models are slightly superior to the quadratic forms when used to forecast Australia's domestic low cost carrier passenger demand.

Table 3. The MAPE values of the GAPAXDE and GARPXSDE models.

GA Function	GAPAXDE MAPE(%)			GARPXSDE MAPE(%)		
	Train	Test	Overall	Train	Test	Overall
Linear	3.34	3.29	3.33	4.65	3.45	4.48
Quadratic	2.38	5.01	3.70	4.02	5.53	4.78

Australia's actual domestic low cost carriers (LCCs) and forecasted enplaned passengers, during the period from Quarter 1 2002 to Quarter 2 2012, are plotted and shown in Figure 2.

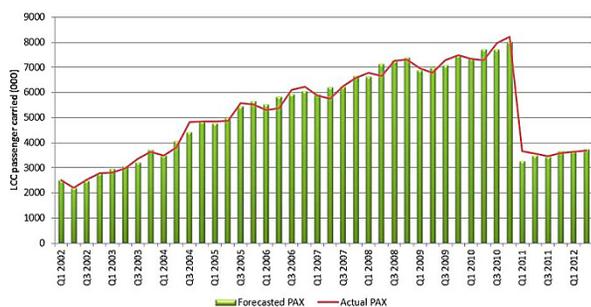


Fig. 2. A comparison of Australia's actual and forecast domestic low cost carriers enplaned passengers (GAPAXDE Model).

Finally, Australia's actual domestic low cost carriers (LCCs) and forecasted revenue passenger kilometres (RPKs), during the period from Quarter 1 2002 to Quarter 2 2012, are plotted and shown in Figure 3.



Fig. 3. A comparison of Australia's actual and forecast domestic low cost carriers RPKs (GARPXSDE Model).

5. Conclusions

This study has developed and empirically examined GA models for forecasting Australia's domestic low cost carriers (LCCs) demand (GAPAXDE and GARPXSDE models). Two mathematical forms, linear, and quadratic, were tested in the study.

Data was divided into two sets, training and testing data set, a 36 training data set is used to estimate the weighting factors of the GA model and a 6 data set was used for testing the robustness of the GA models. The GA parameters used in this study comprised the population size (n): 1000, the generation number: 200, and mutation rate: 0.01. The modelling results have shown that both the linear GAPAXDE and GARPXSDE models are more accurate, reliable and have greater predictive capability as compared to the quadratic form. The overall MAPE of the GAPAXDE and GARPXSDE models are 3.33 per cent and 4.48 per cent, respectively.

References

Abed, S. Y.; Ba-Fail, A. O.; Jasimuddin, S. M. 2001. An econometric analysis of international air travel demand in Saudi Arabia, *Journal of Air Transport Management* 7(3): 143–148. [http://dx.doi.org/10.1016/S0969-6997\(00\)00043-0](http://dx.doi.org/10.1016/S0969-6997(00)00043-0)

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

Akgüngör, A. P.; Doğan, E. 2009. An artificial intelligent approach to traffic accident estimation: model development and application, *Transport* 24(2): 135–142. <http://dx.doi.org/10.3846/1648-4142.2009.24.135-142>

Alekseev, K. P. G.; Seixas, J. M. 2009. A multivariate neural forecasting modeling for air transport – preprocessed by decomposition: a Brazilian application, *Journal of Air Transport Management* 15(5): 212–216. <http://dx.doi.org/10.1016/j.jairtraman.2008.08.008>

Amjadi, M. H.; Nezamabadi-pour, H.; Farsangi, M. M. 2010. Estimation of electricity demand of Iran using two heuristic algorithms, *Energy Conservation and Management* 51(3): 493–497. <http://dx.doi.org/10.1016/j.enconman.2009.10.013>

Azadeh, A.; Saberi, M.; Anvari, M., et al. 2011. An adaptive network based fuzzy inference system-genetic algorithm clustering ensemble algorithm for performance assessment and improvement of conventional power plants, *Expert Systems with Application* 38(3): 2224–2234. <http://dx.doi.org/10.1016/j.eswa.2010.08.010>

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

Baker, D.; Donnet, T. 2012. Regional and remote airports under stress in Australia, *Research in Transportation Business & Management* 4(October): 37–43. <http://dx.doi.org/10.1016/j.rtbm.2012.06.011>

Belobaba, P. P. 2009. Overview of airline economics, markets and demand, in Belobaba, P.; Odoni, A.; Barnhart, C. (Eds.). *The global airline industry*. Chichester, UK: John Wiley & Sons, 47–72. <http://dx.doi.org/10.1002/9780470744734.ch3>

Bhadra, D. 2003. Demand for air travel in the United States: bottom-up econometric estimation and implications for

- forecasts by origin and destination pairs, *Journal of Air Transportation* 8(2):19–56.
- Blinova, T. O. 2007. Analysis of possibility of using neural network to forecast passenger traffic flows in Russia, *Aviation* 11(1): 28–34.
- Carvajal-Rodríguez, A.; Carvajal-Rodríguez, F. 2009. Population extinction in genetic algorithms: application in evolutionary studies, in Omatu, S.; Rocha, M. P.; Bravo, J., et al. (Eds.). *Distributed computing, artificial intelligence, bioinformatics, soft computing, and ambient assisted living*, Lecture Notes in Computer Science Volume 5518. Berlin: Springer, 1028–1034. http://dx.doi.org/10.1007/978-3-642-02481-8_156
- Coelho, R. F.; Méndez Echenagucia, T.; Pugnale, A., et al. 2014. Genetic algorithms for structural design, in Adriessens, S.; Block, P.; Veenendaal, D.; et al. (Eds.). *Shell structures for architecture: form finding and optimization*. Abingdon, UK: Routledge, 290–294.
- Debelle, G. 2009. *Some effects of the global financial crisis on Australian Financial Markets*, Speech to Finance Professionals Forum, Sydney, 31 March 2009, [online], [cited 25 August 2014]. Available from the Internet: <http://www.rba.gov.au/publications/smp/index.html>
- Doganis, R. 2009. *Flying off course: airline economics and marketing*. 4th ed. Abingdon, UK: Routledge.
- Easdown, G.; Wilms, P. 2002. *Ansett: the collapse*. Melbourne: Lothian Books.
- Forsyth, P. 2003. Low-cost carriers in Australia: experiences and impacts, *Journal of Air Transport Management* 9(5): 277–284. [http://dx.doi.org/10.1016/S0969-6997\(03\)00035-8](http://dx.doi.org/10.1016/S0969-6997(03)00035-8)
- Gen, M.; Cheng, R. 1997. *Genetic algorithms and engineering design*. 1st ed. New York: John Wiley & Sons.
- Ghanbari, A.; Kzemi, S. M. R.; Mehmanpazir, F.; et al. 2013. A cooperative ant colony optimization-genetic algorithm approach for construction of energy demand forecasting knowledge-based expert systems, *Knowledge-Based Systems* 39(February): 194–206. <http://dx.doi.org/10.1016/j.knsys.2012.10.017>
- Godinho, P.; Silva, M. 2014. Genetic, memetic and electro-magnetism-like algorithms: applications in marketing, in Moutinho, L.; Bigné, E.; Manrai, A. K. (Eds.). *The Routledge companion to the future of marketing*. Abingdon, UK: Routledge, 393–416.
- Goldberg, D. E. 1989. *Genetic algorithms in search, optimization, and machine learning*. Reading, MA: Addison-Wesley.
- Haldenbilen, S.; Ceylan, H. 2005. Genetic algorithm approach to estimate transport energy demand in Turkey, *Energy Policy* 33(1): 89–98. [http://dx.doi.org/10.1016/S0301-4215\(03\)00202-7](http://dx.doi.org/10.1016/S0301-4215(03)00202-7)
- Hernández-López, M.; Cáceres-Hernández, J. J. 2007. Forecasting tourists' characteristics by a genetic algorithm with a transition matrix, *Tourism Management* 28(1): 290–297. <http://dx.doi.org/10.1016/j.tourman.2005.11.016>
- Holland, J. H. 1975. *Adaption in natural and artificial systems*. Ann Arbor: University of Michigan Press.
- Holloway, S. 2008. *Straight and level: practical airline economics*. 3rd ed. Aldershot, UK: Ashgate Publishing.
- Hong, W. C.; Dong, Y.; Chen, L. Y. et al. 2011. SVR with hybrid chaotic genetic algorithms for tourism demand forecasting, *Applied Soft Computing* 11(2): 1881–1890. <http://dx.doi.org/10.1016/j.asoc.2010.06.003>
- Hu, C. 2002. *Advanced tourism demand forecasting: artificial neural network and Box-Jenkins modelling*: unpublished PhD thesis. Purdue University, USA [online], [cited 6 August 2014]. Available from internet: <http://docs.lib.purdue.edu/dissertations/AAI3104960/>
- Hurley, S.; Moutinho, L.; Witt, S. 1998. Genetic algorithms for tourism marketing, *Annals of Tourism Research* 25(2): 498–514. [http://dx.doi.org/10.1016/S0160-7383\(97\)00099-6](http://dx.doi.org/10.1016/S0160-7383(97)00099-6)
- International Civil Aviation Organization. 2006. *Manual on air traffic forecasting*. Document 8991 AT722/3. 3rd ed. Montreal: ICAO.
- Jirong, G.; Mingcang, Z.; Liuguangyan, J. 2011. Housing price forecasting based on genetic algorithm and support vector machine, *Expert Systems with Applications* 38(4): 3383–3386. <http://dx.doi.org/10.1016/j.eswa.2010.08.123>
- Jones, K. O.; Romil, S. 2004. Selection of an optimum feed profile using genetic algorithms, in Pons, M. N.; Van Impe, J. (Eds.). *Computer applications in biotechnology 2004: a Proceedings volume from the 9th IFAC International Symposium*, Nancy, France, 28–31 March 2004, 577–582.
- KPMG. 2006. *Economic impact study of the Melbourne 2006 Commonwealth Games Post-event analysis*, Office of Commonwealth Games Coordination [online], [cited 15 April 2014]. Available from the Internet: http://www1.dvc.vic.gov.au/ocgc/News/econ_impact_report.pdf
- Kopsch, F. 2012. A demand model for domestic air travel in Sweden, *Journal of Air Transport Management* 20(May): 46–48. <http://dx.doi.org/10.1016/j.jairtraman.2011.11.006>
- Kunt, M. M.; Aghayan, I.; Noii, N. 2011. Prediction for traffic accident severity: comparing the artificial neural network, genetic algorithm, combined genetic algorithm and pattern search methods, *Transport* 26(4): 353–366. <http://dx.doi.org/10.3846/16484142.2011.635465>
- Nolan, J. L. (Ed.). 1996. *Australia business: the portable encyclopedia for doing business with Australia*. San Rafael, CA: World Trade Press.
- Ozturk, H. K.; Ceylan, H.; Canyur, O., E. et al. 2005. Electricity estimation using genetic algorithm approach: a case study of Turkey, *Energy* 30(7): 1003–1012. <http://dx.doi.org/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, in Hall, C. M.; Timothy, D. J.; Duval, D. T. (Eds.). *Safety and security in tourism: relationships, management, and marketing*. Binghamton, NY: The Haworth Hospitality Press, 281–298.
- Sineglazov, V.; Chumachenko, E.; Gorbatyuk, V. 2013. An algorithm for solving the problem of forecasting, *Aviation* 17(1): 9–13. <http://dx.doi.org/10.3846/16487788.2013.777219>
- Sivrikaya, O.; Tunç, E. 2013. Demand forecasting for domestic air transportation in Turkey, *The Open Transportation Journal* 7: 20–26. <http://dx.doi.org/10.2174/1874447820130508001>
- Whyte, R.; Prideaux, B.; Sakata, H. 2012. The evolution of Virgin Australia from a low-cost carrier to a full-service airline – Implications for the tourism industry, in Chen, J. S. (Ed.). *Advances in hospitality and leisure*. Volume 8, Bingley, UK: Emerald Group Publishing, 215–231. [http://dx.doi.org/10.1108/S1745-3542\(2012\)000008015](http://dx.doi.org/10.1108/S1745-3542(2012)000008015)