AIR PASSENGER AND AIR CARGO DEMAND FORECASTING: APPLYING ARTIFICIAL NEURAL NETWORKS TO EVALUATING INPUT VARIABLES

Shih-Yao Kuo, Dept. of Aviation Service Management, Aletheia Univ., Taiwan.

Shu-Chuan Chen, Dept. of Aviation Service Management, Aletheia Univ., Taiwan.

ABSTRACT

The aviation industry relies strongly on air traffic demand forecasting to develop operating strategies, including those related to destinations and routes, fleet planning and routing, and human resources. Time series analysis, gravity models, grey theory and artificial neural networks are common tools for air traffic demand forecasting. This study employed artificial neural networks to forecast the air passenger and air cargo demand from Japan to Taiwan. All input variables were collected based on a literature survey, market analysis and preliminary evaluation. The factors which influence air passenger and air cargo demand were identified, evaluated and analyzed in detail by applying artificial neural networks. The results reveal that some factors influence both passenger and cargo demand, and the others only one of them. The employed population in Japan and per capita income (PCI) in Taiwan influence both air passenger and air cargo volume. Flight movement from Tokyo (NRT) to Taipei (TPE), PCI in Taiwan, and foreign exchange rate are the three most important factors for air passenger volume, while the economic growth rate in Taiwan is the most important factor for air cargo volume. By using a neural network, a novel forecasting model that considers actual air passenger and air cargo demand was established, and the results show that it is an accurate tool to forecast air traffic demand. In addition, the valuable data that can be obtained from the model may be a good reference for air carriers and government authorities.
Keywords: artificial neural networks, air passenger, air cargo, forecast

1. INTORDUCTION

The aviation industry relies strongly on air traffic demand forecasting to develop operating strategies with regard to destinations and routes, fleet planning and routing, and human resources. Forecasting can be achieved by qualitative or quantitative approaches, with advances in technology meaning that the latter can produce more objective and reliable results. Time series analysis, gravity models, grey theory and artificial neural networks are some of the commonly used tools for air traffic demand forecasting.

Most of the publicly available forecasting figures in the aviation industry are based on the time-series and econometric models, and most often they represent global or regional data (Airbus, 2009; Boeing, 2009). However, the benefit to airlines of such figures is limited, as there is a lack of data for specific routes or markets. In addition, most of the academic literature focuses on forecasting figures for passengers, cargo or traffic (combining air passenger and air cargo) individually. To date, the factors which influence both passenger and cargo demand together have not been studied in detail by a quantitative approach.

In this study, the factors which influence air passenger and cargo demand have been identified, evaluated and analyzed in detail by applying artificial neural networks. As visitors from Japan are Taiwan's primary inbound market segment, and Japan is one of Taiwan's major trading partners, the Japan-Taiwan country-pair is chosen for empirical analysis. A novel forecasting model that considers the actual air passenger and air cargo markets is established in this work, with results that will benefit both air carriers and government authorities.

2. AIR TRAFFIC DEMAND FORECASTING

2.1 Air passengers

A comprehensive review of research on air passenger demand was presented by Jorge-Calderon (1997). A demand model for scheduled airline services was established for the entire network of international European routes in 1989. The estimated model included variables describing both the geo-economic characteristics of the area where transportation took place and the patterns of airline service, as determined by flight frequency, plane size and prices. Brons et al. (2002) collected 37 studies and 204 observations, and indicated that air passenger demand is largely determined by the spending capacity of customers. Hsu and Wen (1998) used grey theory to develop time series GM(1,1) models for forecasting total passenger and 10 country-pair passenger traffic flows in the Asia-Pacific market. In addition, Hsu and Wen (2000) later developed a series of models capable of forecasting airline city-pair passenger traffic, designing a network of airline routes and determining flight frequencies on individual routes were developed by applying grey theory and multi-objective programming.

More recently, Grosche et al. (2007) presented two gravity models for the estimation of air passenger volume between city-pairs. The models include variables describing the general economic activity and geographical characteristics of city-pairs instead of variables describing air service characteristics. Booking data of flights between Germany and 28 European countries were used for calibration, and both models show a good fit to the observed data. Kuo et al. (2010) employed artificial neural networks (ANN) to establish a mathematical model with multiple inputs and multiple outputs, and their results indicate that this model may accurately forecast the air transport demand for a network of routes.

2.2 Air cargo

The myriad of factors that influence the choice of airport by freighter-operating airlines were identified and evaluated by Gardiner et al. (2005). A number of factors such as night curfews, freight forwarders and airport charges were found to be influential, and these were examined against a number of dependent variables such as the airline's home region and operational patterns in order to identify key variations. Ohashi et al. (2005) used air cargo traffic flow data to identify the critical factors influencing air cargo transshipment route choice decisions. The empirical results showed that the choice of the air cargo transshipment hub is more sensitive to time cost than the monetary costs such as landing fees and line-haul price.

Zhang and Zhang (2002) employed a multi-market oligopoly model to examine the effect of cargo liberalization on competition between all-cargo carriers and mixed passenger/cargo carriers. They suggested that the separation of air cargo and passenger rights might be fraught with difficulty in Asia due to the characteristics of its air cargo market, in which most passenger carriers have substantial cargo businesses and operate ''combi'' fleets. Hong Kong air cargo in the context of both China and the Asia-Pacific region was discussed by Zhang (2003), and an oligopoly model was also developed to investigate the effect of an air cargo alliance on competition in passenger markets by Zhang et al. (2004). Lin and Chen (2003) analyzed government documents and interviewed the air cargo carriers and airlines that serve the Taiwan-China air cargo market. The information thus obtained enabled them to tabulate the trade between Taiwan and China, estimate the airport-to-airport air cargo demand and calibrate the international and domestic freight tariffs. They also developed a mathematical model and a branch-and-bound algorithm, and the results showed that at least two transit airports are economically necessary for a Taiwan-China air link. Shanghai and Xiamen were always the top two transit airport choices in their model, and the third airport would be Changsha if three air-links were chosen. These links are different from the top three passenger transit airports, which were Fuzhou, Xiamen and Shanghai.

Matsumoto (2004) examined international urban systems from the standpoint of international air traffic flows and analyzed the patterns of international air passenger and cargo flows within and among the Asian, European and American regions. After evaluating the international air network structures, the degree of 'hub-ness' for prospective hub cities was clarified by a basic gravity model composed of gross domestic product (GDP), population and distance introducing city-dummy variables. This earlier work was extended by Matsumoto (2007), which separated international air network structures for passengers and cargo to more precisely describe airport traffic volumes. The results revealed that many cities are strengthening their position as international air transportation hubs, especially: Tokyo, Hong Kong, Singapore, London, Paris, Frankfurt, Amsterdam, New York, and Miami. Moreover, the air traffic density of three cities in particular, Seoul, Hong Kong, and Amsterdam, is growing at an extraordinary rate.

2.3 Forecasting by artificial neural networks

Neural networks were first applied to forecast tourism demand by Law and Au (1999) to predict Japanese demand for travel to Hong Kong. The independent variables in their work included service price in Hong Kong relative to Japan, foreign exchange rate, population in Japan, marketing expenses to promote Hong Kong's tourism industry, real gross domestic expenditure per person in Japan, and average hotel rate in Hong Kong. The results showed that the neural network model outperformed multiple regression, moving average, and exponent smoothing approaches. Law (2000) extended the applicability of neural networks in tourism demand forecasting by incorporating the back-propagation learning process into nonlinearly separable tourism demand data. The empirical results indicated that a back-propagation neural network outperforms regression models, time-series models and feed-forward neural networks in terms of forecasting accuracy.

Cho (2003) also concluded that an ANN approach outperforms the exponential smoothing and auto-regressive integrated moving average (ARIMA) models with regard to tourism demand forecasting. Wei and Yang (1999) built an artificial neural network model to forecast the total transshipment container throughput of Kaohsiung port, using the following variables from Taiwanese economic data: gross domestic product, economy growth rate, industry index of product, average national income, wholesale price index, and average gross national product.

The literature survey shows that population, per capita income (PCI), gross domestic product (GDP), and consumer price index (CPI) are the most commonly used independent variables in previous studies. In addition, time series analysis, gravity models, grey theory and artificial neural networks are the tools that are usually selected to forecast air traffic demand, and relatively few studies have been conducted using artificial neural networks. While there is no doubt that the air passenger and air cargo demand are different, the factors that influence both of these have not yet discussed together and in detail by quantitative approach.

3. METHOD AND DATA

3.1 Back-propagation neural networks

A neural network consists of an input layer, an output layer, and usually one or more hidden layers (Masters, 1993; Haykin, 1999). Each of these layers contains nodes, and these are connected to other nodes at adjacent layers. Fig. 1 illustrates a neural network with three layers. Each node in a neural network is a processing unit. The output from a given neuron is calculated by applying a transfer function to a weighted summation of its input to give an output. In this study, the summation function is the weighted summation.

$$I_j = \sum_i W_{ij} \times X_i \tag{1}$$

The activity function represents the output of the summation function.

$$net_j = I_j \tag{2}$$

The nonlinear transfer function has a sigmoid shape.

$$Y_j = \frac{1}{1 + exp^{-net_j}} \tag{3}$$

Because the output range of the sigmoid function is between 0 and 1, the mapping of data is more suitable. The following simple equation is employed because it is not necessary to adjust the data pattern.

$$X_{new} = \frac{\left(X_{old} - X_{min}\right)}{\left(X_{max} - X_{min}\right)} \times \left(D_{max} - D_{min}\right) + D_{min} \tag{4}$$

where X_{max} and X_{min} are the maximum and minimum of the variables. D_{max} (=0.8) and D_{min} (=0.1) are the specified maximum and minimum.

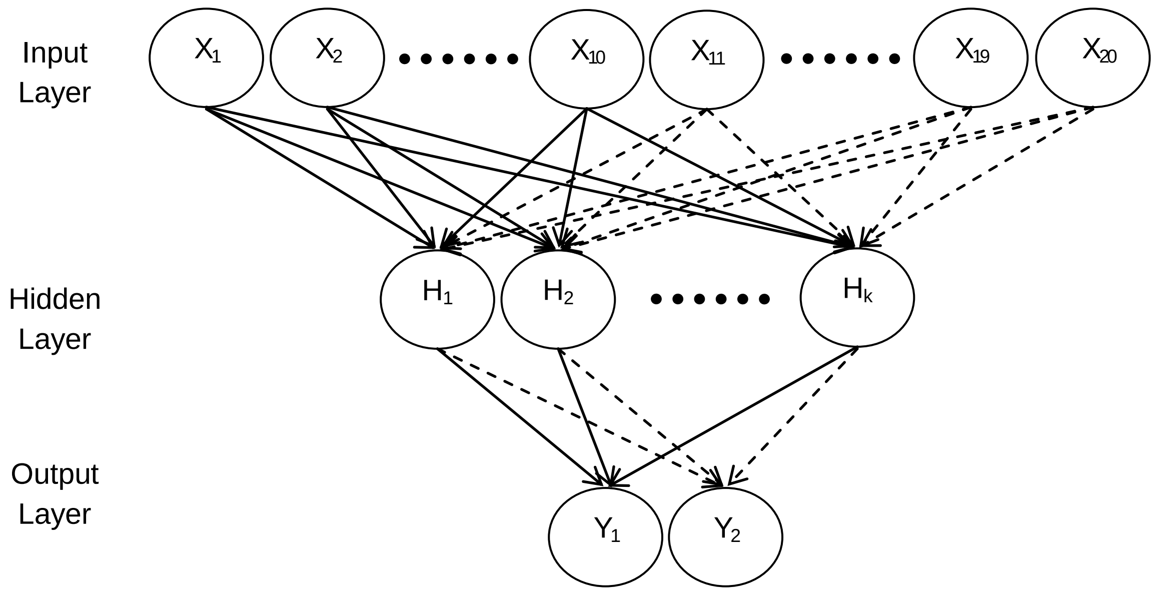

Figure 1 – Artificial neural network

Nodes in the input layer represent independent variables of the problem, including the appropriate social and economic parameters. The hidden layer is used to add an internal representation of nonlinear data. In general, the number of nodes in the hidden layer is determined by the arithmetic or geometric mean of the nodes in the input and output layers. In this research, the arithmetic mean was used for fast convergence and stable performance. The output of a neural network is the solution to a problem. In this study, the numeric values from the output nodes were used to represent air passenger and air cargo volumes. The energy function is a verification function which determines if the network energy has converged to its minimum. Whenever the energy function approaches zero, the network approaches its optimum solution.

$$E = \frac{1}{2}\sum_{j=1}^{n}\left(T_j - Y_j\right)^2 \tag{5}$$

where T_j is the actual observation, Y_j is the predicted value, and n is the number of predictions. The gradient steepest descent method was employed to find the minimum of the energy function.

Back-propagation network (BPN) is one of the most commonly used supervised ANN models. BPN is a layered feed-forward supervised network which is suitable for use in the air passenger and air cargo forecasting models presented in this study. A simple MATLAB computer program was developed for the neural network model, as shown in Fig. 1. In this study, seven widely used criteria: *MAE* (=*MAD*), *MAPE*, *SSE*, *MSE*, *RMS*E, *RMSPE*, and *r* (Law and Au, 1999), were used to evaluate the forecasting performance of neural networks. The measures of forecasting accuracy are listed as follows:

Mean Absolute Error: $$MAE = MAD = \frac{1}{n}\sum_{j=1}^{n}\left|T_j - Y_j\right| \tag{6}$$

Mean Absolute Percentage Error: $$MAPE = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{T_j - Y_j}{T_j}\right| \times 100\% \tag{7}$$

Sum of Square Error: $$SSE = \sum_{j=1}^{n}\left(T_j - Y_j\right)^2 \tag{8}$$

Mean Square Error: $$MSE = \frac{1}{n}SSE = \frac{1}{n}\sum_{j=1}^{n}\left(T_j - Y_j\right)^2 \tag{9}$$

Root Mean Square Error: $$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(T_j - Y_j\right)^2} \tag{10}$$

Root Mean Square Percentage Error: $$RMSPE = \frac{1}{n}\sqrt{\sum_{j=1}^{n}\left(\frac{T_j - Y_j}{T_j}\right)^2} \times 100\% \tag{11}$$

Normalized Correlation Coefficient: $$r = \frac{\sum_{j=1}^{n} T_j \times Y_j}{\sqrt{\sum_{j=1}^{n} T_j^{\,2} \times \sum_{j=1}^{n} Y_j^{\,2}}} \tag{12}$$

3.2 Data collection and variables used

Secondary sources of data were used in this research. All input variables were collected through a literature survey, market analysis and preliminary evaluation. The preliminary selection of data for the neural network was based on representative social and economic parameters, data availability, the reliability of data sources, and the measurability of variables in the modelling process (Law and Au, 1999).

Gately (1996) indicated that the transfer function can handle the collinear variables in artificial neural networks. Wei and Yang (1999) also reported that the forecasting performance of artificial neural networks will not be influenced by input variables that are correlated with each other. Therefore, the following 10 variables were chosen for the candidate input variables for forecasting air passenger demand from Japan to Taiwan: population in Japan, employed population in Japan, PCI in Japan, GDP in Japan, GNP in Japan, economic growth rate in Japan, foreign exchange rate, flight movement from Tokyo (NRT) to Taipei (TPE), CPI in Taiwan, and average hotel rate in Taiwan. The following 10 variables were chosen for the candidate input variables for forecasting air cargo demand from Japan to Taiwan: population in Taiwan, employed population in Taiwan, PCI in Taiwan, GDP in Taiwan, GNP in Taiwan, economic growth rate in Taiwan, import price index in Taiwan, listed companies in Tokyo, gross import amount from Japan to Taiwan, and industrial product in Japan.

Table 1 – Candidate input variables and output variable for air passenger demand

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	Y_1
Year	Population in Japan (1000)	Employed population In Japan (1000)	PCI in Japan (1000 ¥)	GDP in Japan (M $)	GNP in Japan (M $)	Economic growth rate in Japan (%)	Foreign exchange rate (¥/NT$)	Flight movement from Tokyo to Taipei	CPI in Taiwan (%)	Average hotel rate in Taiwan (NT$/Night)	Air passenger demand from Japan to Taiwan
1999	126686	64620	2915	4368731	3941600	-0.10	0.3044	8449	94.90	2897	826222
2000	126926	64460	2929	4667471	3871200	2.90	0.2944	8565	96.09	2947	916301
2001	127291	64120	2840	4095083	3673200	0.20	0.2667	8765	96.08	2951	976750
2002	127486	63300	2791	3918333	3986200	0.30	0.2930	10080	95.89	2907	998497
2003	127694	63160	2804	4229115	4302800	1.40	0.3179	9012	95.62	2763	657053
2004	127787	63290	2849	4605917	4695400	2.70	0.3110	9702	97.17	2913	887311
2005	127768	63560	2865	4552197	4662200	1.90	0.2798	9931	99.41	2987	1124334
2006	127770	63820	2924	4376005	4486500	2.00	0.2741	11456	100.00	3124	1161489
2007	127771	64120	2934	4370749	4526500	2.40	0.2897	11400	101.80	3220	1166380
2008	127663	63850	2986	4376005	5073100	-0.70	0.3641	11387	105.39	3214	1086691

Table 1, containing the relevant data for forecasting air passenger demand from Japan to Taiwan in the period of 1999 to 2008, was set up based on public sources. Table 2 shows the relevant data for forecasting air cargo from Japan to Taiwan in the period of 1999 to 2008. Most of the data in Taiwan were collected from the Directorate General of Budget, Accounting and Statistics of Executive Yuan, Taiwan (2009). Flight movement from Tokyo to Taipei, air passengers from Japan to Taiwan, air cargo from Japan to Taiwan, and average hotel rate in Taiwan were collected from the Ministry of Transportation and Communications of Executive Yuan (2009). The amount for gross imports from Japan to Taiwan was collected from the Ministry of Economic Affairs of Executive Yuan (2009). The foreign exchange rate data was collected from the Central Bank (2009). The data for Japan were collected from Ministry of Internal Affairs and Communications, Japan (2009).

Table 2 – Candidate input variables and output variable for air cargo demand

	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	Y_2
Year	Population in Taiwan (1000)	Employed population in Taiwan (1000)	PCI in Taiwan ($)	GDP in Taiwan (M $)	GNP in Taiwan (M $)	Economic growth rate in Taiwan (%)	Import price index in Taiwan (%)	Listed companies in Tokyo	Gross import amount from Japan to Taiwan (1000 $)	Industrial product in Japan (B ¥)	Air cargo demand from Japan to Taiwan (Kg)
1999	22092	9385	12324	298757	301562	5.75	75.77	1892	30637138	110125	84782650
2000	22277	9491	13090	321230	325698	5.77	79.28	2055	38622073	111439	95121480
2001	22406	9383	11692	291694	297374	-2.17	78.29	2103	25933484	104039	70213408
2002	22521	9454	11914	297668	304680	4.64	78.60	2119	27363325	101272	84839123
2003	22605	9573	12242	305624	315179	3.50	82.64	2174	32719525	102757	97747098
2004	22689	9786	13252	331007	342137	6.15	89.72	2276	43717802	105410	123062797
2005	22770	9942	14075	355958	364997	4.16	91.90	2323	46053319	107877	107541270
2006	22877	10111	14455	366357	375939	4.80	100.00	2391	46284409	108603	114089127
2007	22958	10294	15122	384768	394901	5.70	108.95	2389	45936862	108601	108026177
2008	23037	10403	15276	392050	402572	0.12	118.58	2374	46508013	108578	87391641

The number of candidate input variables were assigned to be the same for both air passenger and air cargo demand. The first six independent variables stand for the variables in Japan for air passenger demand and the ones in Taiwan for air cargo demand. For example, X_1 is population in Japan and X_{11} is population in Taiwan. The former may be the driving force to influence air passenger demand from Japan to Taiwan, while the later may be the driving force to influence air cargo demand from Japan to Taiwan. It is well-known that economic growth drives air traffic demand. Therefore, X_6 (Economic growth rate in Japan) and X_{16} (Economic growth rate in Taiwan) were employed. In addition, population and employed population were simultaneously used, as were GDP and GNP. These collinear variables were used to demonstrate the forecasting performance of artificial neural networks.

X_7 (Foreign exchange rate), X_9 (CPI), and X_{10} (Average hotel rate) were similar with the input variables used by Law and Au (1999). Flight movement from Tokyo (NRT) to Taipei (TPE) was also employed to represent the most important route connecting the two capitals. After market analysis and preliminary evaluation, X_{17} (Import price index in Taiwan), X_{18} (Listed companies in Tokyo), X_{19} (Gross import amount from Japan to Taiwan), and X_{20} (Industrial product in Japan) were also proposed as the candidate input variables for air cargo. Gross import amount from Japan to Taiwan includes air and marine freight. The above 20 factors which may influence air passenger and air cargo demand were identified, evaluated and analyzed in detail by applying artificial neural networks, as follows.

4. RESULTS AND DISCUSSION

The empirical results are discussed in the following four sections. First is air passenger demand forecasting, followed by air cargo demand forecasting. Next, the candidate variables have been deleted and added to verify their influences. Finally, the hybrid model in which air passenger and air cargo are the two dependent variables is presented.

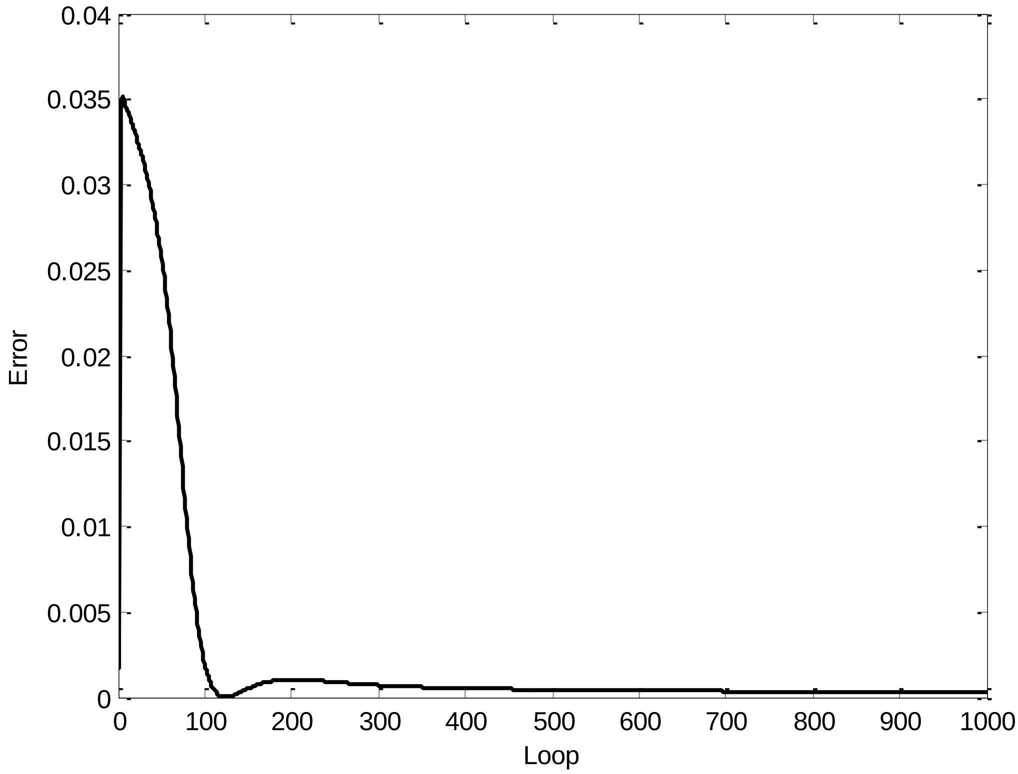

Figure 2 – Energy function vs. iteration loop

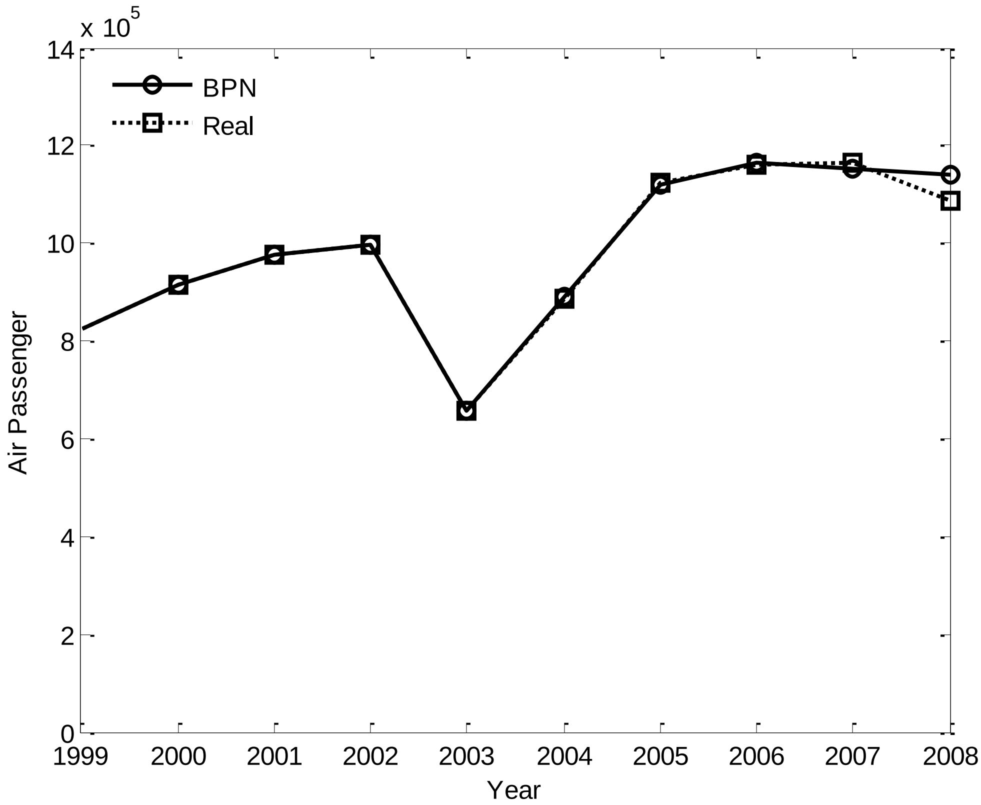

Figure 3 – Air passenger demand forecasting

4.1 Air passenger demand forecasting

This study adapts the data from Table 1 for network training and validation. Considering the requirements of the aviation industry, two years may be a reasonable lead time for developing an operating strategy. Among the 10 years' sample entries, the first eight rows from 1998 to 2006 are selected to form the training set, while the rest are used for forecasting accuracy validation. In general, a large sample size is often suggested for training in order to obtain sufficient learning. However, the literature offers little guidance in either selecting the samples or their sizes. Kuo et al. (2010) employed an ANN model to forecast the air transport demand in a network of routes network, and their empirical results for eight different cases reveal that a small sample size (SSS) is tolerable and acceptable in terms of forecasting performance for ANN model.

The 10 candidate variables in Table 1 are all used for the 10 nodes (X_1, X_2……X_{10}) in the input layer. The output variable (Y_1) used in this research is air passengers from Japan to Taiwan. Fig. 1 (solid lines) shows a neural network with a hidden layer of six nodes, which is used to determine air passenger demand. In this study, six nodes are used for fast convergence and stable performance. The energy function vs. iteration loop is shown in Fig. 2. It is seen that the present model can achieve a convergent result after the training phase. Fig. 3 provides a graphical presentation of the empirical findings with the BPN air passenger demand forecasting model. The dotted line is the actual observation and the solid line is the predicted value. They are close to each other in spite of the air passenger volume declining due to the SARS epidemic in 2003.

Table 3 shows the experimental results of three training year models and seven criteria. Because their values vary case by case and become less referable, the following four criteria are not recommended to evaluate forecasting accuracy: Mean Absolute Error (*MAE*, *MAD*), Sum of Square Error (*SSE*), Mean Square Error (*MSE*), and Root Mean Square Error (*RMSE*). In addition, Normalized Correlation Coefficient (*r*) would malfunction if the number of predictions is small. Because the differences between the actual observations and the predicted values are squared in calculating Root Mean Square Percentage Error (*RMSPE*), the differences would increase if they are more uneven. Therefore, Mean Absolute Percentage Error (*MAPE*) is the recommended criterion to evaluate forecasting accuracy. The scale developed by Lewis (1982) is used to evaluate forecasting performance. In this scale, any forecast with *MAPE* less than 10% is considered highly accurate, 10-20% is good, 20-50% reasonable, and greater than 50% inaccurate. All forecasts of the present novel BPN model fall in the highly accurate range. The smallest *MAPE* (=3.07%) is achieved for the eight-year training model, and thus this is used in the following analyses.

Table 3 – Experimental results of air passenger demand forecasting

Training	Eight years (1999-2006)	Seven years (2000-2006)	Six years (2001-2006)
Loop	3350	4110	3012
MAPE (%)	3.07	3.32	3.30
RMSPE (%)	2.58	3.05	2.79
r	0.9995	0.9995	0.9995
MAE(*MAD*)	33.81×10^3	36.30×10^3	36.36×10^3
SSE	3.16×10^9	4.39×10^9	3.71×10^9
MSE	1.58×10^9	2.19×10^9	1.86×10^9
RMSE	39.76×10^3	46.84×10^3	43.07×10^3

4.2 Air cargo demand forecasting

The 10 candidate variables in Table 2 are used for the 10 nodes (X_{11}, X_{12}……X_{20}) in the input layer. The output variable (Y_2) used in this research is air cargo demand from Japan to Taiwan. Fig. 1 (dashed lines) shows the neural network, with a hidden layer of six nodes. Fig. 4 provides a graphical presentation of the empirical findings with the BPN air cargo demand forecasting model. Differences between the solid line (predicted values) and the dotted line (actual observations) are observed. It is noteworthy that the air cargo demand from Japan to Taiwan fell dramatically by 19% due to the economic crisis that began in 2008. However, the forecast of the present BPN model with *MAPE* = 14.56% still falls in the good range. Notably, the reversal of air cargo demand is indicated in this BPN model.

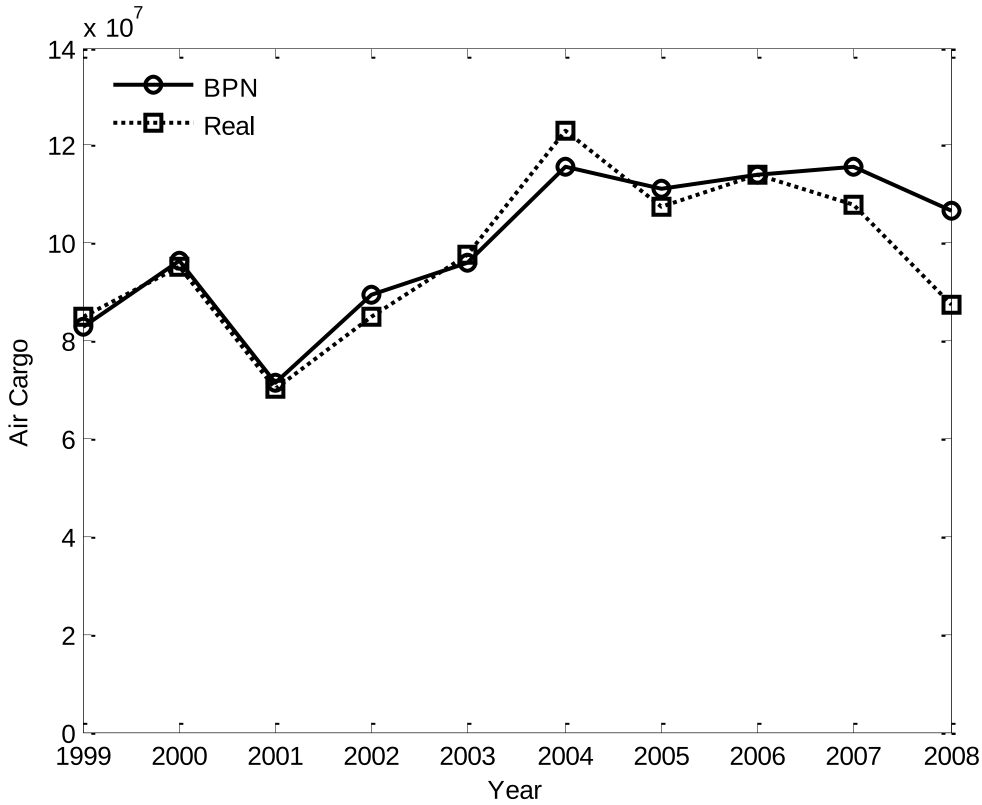

Figure 4 – Air cargo demand forecasting

4.3 The effects of deleting and adding input variables

In this section, the 10 candidate variables in Table 1 are deleted separately. If the deleting of a candidate variable reduces the *MAPE* value, then it interferes with air passenger demand forecasting and thus should be excluded from the BPN model. After deleting X_6 (Economic growth rate in Japan), X_9 (CPI in Taiwan), and X_{10} (Average hotel rate in Taiwan), the BPN model can more accurately forecast with *MAPE* = 0.39%, as shown in Table 4. The cross symbol indicates the input variables are included in the BPN model.

The candidate variables in Table 2 are then added step by step into the model. If the adding of a candidate variable makes *MAPE* value lower, then it is a necessary influencing factor for air passenger demand forecasting, and thus should be included in the BPN model. After adding X_{13} (PCI in Taiwan), the BPN model can more accurately forecast air passenger demand with an extremely low *MAPE* = 0.34%, as shown in Table 4. The *MAPE* values that appear in Table 4 with lines through them lines are to distinguish the worse approaches from the better ones.

It is interesting to note that X_9 (CPI in Taiwan) and X_{10} (Average hotel rate in Taiwan) should be excluded to establish the more accurate forecasting model for air passenger demand from Japan to Taiwan, including business and leisure passengers. However, the service price (relative CPI) and average hotel rate were used in the ANN model to forecast tourism demand (Law and Au, 1999; Law, 2000). In these two earlier works, the above procedure was not presented and their *MAPE* values were higher.

Table 4 –The effects of deleting and adding candidate variables on air passenger forecasting

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	*MAPE* (%)
x	x	x	x	x	x	x	x	x	x											**3.07**
	x	x	x	x	x	x	x	x	x											~~3.42~~
x		x	x	x	x	x	x	x	x											~~3.46~~
x	x		x	x	x	x	x	x	x											~~3.31~~
x	x	x		x	x	x	x	x	x											~~3.15~~
x	x	x	x		x	x	x	x	x											~~3.28~~
x	x	x	x	x		x	x	x	x											2.14
x	x	x	x	x	x		x	x	x											~~3.47~~
x	x	x	x	x	x	x		x	x											~~5.20~~
x	x	x	x	x	x	x	x		x											1.20
x	x	x	x	x	x	x	x	x												2.38
x	x	x	x	x		x	x		x											0.60
x	x	x	x	x		x	x	x												1.10
x	x	x	x	x	x	x	x													1.11
x	x	x	x	x		x	x													**0.39**
x	x	x	x	x		x	x			x										~~0.66~~
x	x	x	x	x		x	x				x									~~0.58~~
x	x	x	x	x		x	x					x								**0.34**
x	x	x	x	x		x	x						x							~~0.46~~
x	x	x	x	x		x	x							x						~~0.55~~
x	x	x	x	x		x	x								x					~~16.75~~
x	x	x	x	x		x	x									x				~~0.81~~
x	x	x	x	x		x	x										x			~~0.50~~
x	x	x	x	x		x	x											x		~~0.42~~
x	x	x	x	x		x	x												x	~~1.65~~

Note: The cross symbol indicates the input variables are included in the BPN model.

The same procedure may be used in air cargo demand forecasting, as shown in Table 5. After deleting X_{20} (Industrial product in Japan) and adding X_2 (Employed population in Japan), the BPN model can more accurately forecast air cargo demand with *MAPE* = 7.74%. The forecast of the present novel BPN model falls in the highly accurate range.

As shown in Fig. 5, the variables inside the left circle are the factors that influence air passenger volume. The variables inside the right circle are the factors that influence air cargo volume. The variables in the overlap between the two circles are underlined to indicate that they influence both air passenger and air cargo volume.

X_1 (Population in Japan), X_3 (PCI in Japan), X_4 (GDP in Japan), X_5 (GNP in Japan), X_7 (Foreign exchange rate), and X_8 (Flight movement from Tokyo to Taipei) only influence air passenger volume. X_{11} (Population in Taiwan), X_{12} (Employed population in Taiwan), X_{14} (GDP in Taiwan), X_{15} (GNP in Taiwan), X_{16} (Economic growth rate in Taiwan), X_{17} (Import price index in Taiwan), X_{18} (Listed companies in Tokyo), and X_{19} (Gross import amount from Japan to Taiwan) only influence air cargo volume. X_2 (Employed population in Japan) and X_{13} (PCI in Taiwan) influence both of air passenger and air cargo volumes.

The candidate variables outside these two circles are the inappropriate independent variables, and thus should be excluded. X_6 (Economic growth rate in Japan), X_9 (CPI in Taiwan), X_{10} (Average hotel rate in Taiwan), and X_{20} (Industrial product in Japan) are the interferences in the BPN model. These variables appear in Fig. 5 with lines through them. Although Japan is Taiwan's largest source of imports, Taiwan is not Japan's primary export market. It is thus seen that X_6 (Economic growth rate in Japan) and X_{20} (Industrial product in Japan) are not the appropriate input variables in the BPN model. Since PCI in Japan is much higher than PCI in Taiwan and business passengers are less sensitive to price, X_9 (CPI in Taiwan) and X_{10} (Average hotel rate in Taiwan) do not moderate the number of air passengers from Japan to Taiwan.

Table 5 –The effects of deleting and adding candidate variables on air cargo forecasting

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	X_{20}	*MAPE (%)*
										x	x	x	x	x	x	x	x	x	x	**14.56**
											x	x	x	x	x	x	x	x	x	~~16.75~~
										x		x	x	x	x	x	x	x	x	~~15.56~~
										x	x		x	x	x	x	x	x	x	~~15.18~~
										x	x	x		x	x	x	x	x	x	~~15.45~~
										x	x	x	x		x	x	x	x	x	~~14.66~~
										x	x	x	x	x		x	x	x	x	~~17.38~~
										x	x	x	x	x	x		x	x	x	~~15.12~~
										x	x	x	x	x	x	x		x	x	~~16.12~~
										x	x	x	x	x	x	x	x		x	~~15.00~~
										x	x	x	x	x	x	x	x	x		**12.39**
x										x	x	x	x	x	x	x	x	x		~~12.41~~
	x									x	x	x	x	x	x	x	x	x		**7.74**
		x								x	x	x	x	x	x	x	x	x		~~14.08~~
			x							x	x	x	x	x	x	x	x	x		~~23.48~~
				x						x	x	x	x	x	x	x	x	x		~~16.97~~
					x					x	x	x	x	x	x	x	x	x		~~24.57~~
						x				x	x	x	x	x	x	x	x	x		~~17.34~~
							x			x	x	x	x	x	x	x	x	x		~~17.78~~
								x		x	x	x	x	x	x	x	x	x		~~14.27~~
									x	x	x	x	x	x	x	x	x	x		~~17.08~~

Note: The cross symbol indicates the input variables are included in the BPN model.

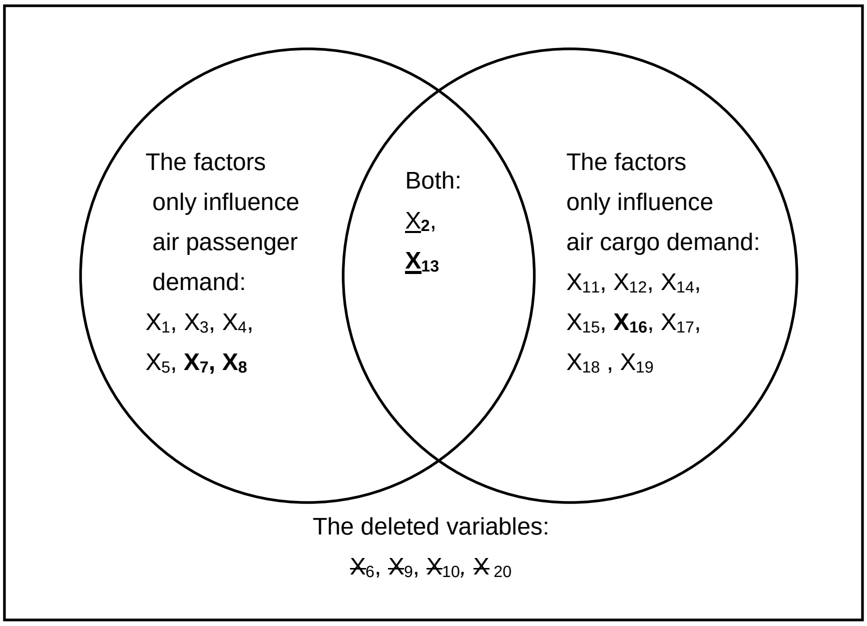

Figure 5 –The factors that influence air passenger and/or air cargo

4.4 Hybrid demand forecasting

If the 16-9-2 BPN model is used for forecasting air passenger and air cargo demand simultaneously, the forecasting accuracy is decreased. It seems that *MAPE* values for air passenger and air cargo demand will increase considerably, as shown in Table 6. The factors in the left half-moon in Fig. 5, namely X_1, X_3, X_4, X_5, X_7, and X_8, show no influence on air cargo demand. Instead, they become interferences in the present hybrid model. The same is true for the factors in the right half-moon in Fig. 5 for air passenger demand. Therefore, the forecasting accuracy is decreased.

Table 6 –The effects of modelling

	Single output						Hybrid	
Output	Passenger			Cargo			Passenger	Cargo
Input	10p	7p	7p+1c	10c	9c	9c+1p	7p+9c	
Model	10-6-1	7-4-1	8-5-1	10-6-1	9-5-1	10-6-1	16-9-2	
MAPE (%)	3.07	0.39	0.34	14.56	12.39	7.74	6.38	15.54

A contribution graph (Gately, 1996) is also used to study the influence of a processing unit in the input layer. The contribution of an input variable is the sum of absolute values of the weighting connected to every node in the hidden layer ($\sum_{j=1}^{k}\left|W_{ij}\right|$). To evaluate the influences of input variables, the scale developed by Gately (1996) is used. Any input variable with a contribution of less than 2 has weak influence, while one with a contribution greater than 5 has a significant influence.

Table 7 shows the contributions of input variables for three models: air passenger, air cargo, and hybrid (forecasting air passenger and air cargo simultaneously). X_8 (Flight movement from Tokyo to Taipei), X_{13} (PCI in Taiwan), and X_7 (Foreign exchange rate) are the three most important factors for air passenger demand forecasting. X_{16} (Economic growth rate in Taiwan) is the most important factor for air cargo demand forecasting. These important input variables are in boldface type in Table 7 and Fig. 5. In addition, the variables influencing both air passenger and air cargo volume are underlined in Table 7 and Fig. 5. The interferences appear in Table 7 and Fig. 5 with lines thorough them.

Table 7 – The contributions of input variables

Output		Passenger	Cargo	Hybrid
Input		7p+1c	9c+1p	7p+9c
BPN model		8-5-1	10-6-1	16-9-2
X_1	Population in Japan	1.82	-	3.44
X_2	Employed population in Japan	3.94	-	3.88
X_3	PCI in Japan	3.20	-	3.34
X_4	GDP in Japan	2.69	-	**5.22**
X_5	GNP in Japan	3.55	-	4.71
X_6	~~Economic growth rate in Japan~~	-	-	-
X_7	**Foreign exchange rate**	**5.63**	-	**7.42**
X_8	**Flight movement from Tokyo to Taipei**	**8.27**	-	4.78
X_9	~~CPI in Taiwan~~	-	-	-
X_{10}	~~Average hotel rate in Taiwan~~	-	-	-
X_{11}	Population in Taiwan	-	2.94	2.29
X_{12}	Employed population in Taiwan	-	3.76	2.79
X_{13}	**PCI in Taiwan**	**6.30**	3.39	4.31
X_{14}	GDP in Taiwan	-	2.16	**5.49**
X_{15}	GNP in Taiwan	-	3.98	4.68
X_{16}	**Economic growth rate in Taiwan**	-	**5.23**	**5.03**
X_{17}	Import Price Index in Taiwan	-	2.31	4.49
X_{18}	Listed companies in Tokyo	-	2.09	4.50
X_{19}	Gross import amount from Japan to Taiwan	-	4.79	3.94
X_{20}	~~Industrial product in Japan~~	-	-	-

It is noteworthy that the results from Tables 4 and 5 obtained by deleting and adding candidate variables are similar to that from Table 7 attained with the use of a contribution graph. For example, if X_8 (Flight movement from Tokyo to Taipei) is excluded in the BPN model, forecasting accuracy would decrease most severely, as shown in Table 4 (MAPE=5.20%). Also, its contribution (=8.27) is highest, as shown in Table 7. However, few conclusions can be drawn by comparing the hybrid model (forecasting air passenger and air cargo demand simultaneously) and the single output models (forecasting air passenger or air cargo demand individually).

It seems that these collinear variables (population and employed population, GDP and GNP) will not cause any difficulty in the present back-propagation neural network forecasting model. The weighting of these collinear variables will be adjusted automatically in the training phase. If only population and GDP are employed, forecasting accuracy would be decreased considerably and their contributions would not be greater than 5.

To reflect air traffic volume, workload unit (WLU) is used as a traffic measure combining air passenger and air cargo demand. A WLU is equivalent to one terminal passenger or a hundred kilograms of cargo handled (Doganis, 1992; Matsumoto, 2004). Comparing Figures 3 and 4, it is interesting to note that the magnitude of air cargo demand is almost 100 times that air passenger demand. The weight of one passenger with baggage is almost equal to 100 kilograms. Air carriers develop both passenger and cargo transportation with equal emphasis in the Japan-Taiwan country-pair. However, air cargo volume may not be proportional to air passenger volume. For example, while air passenger volume fell dramatically due to the SARS epidemic in 2003, air cargo volume continued to rise. If WLU is used as the dependent variable combining air passenger and air cargo demand, the present BPN model could not forecast air traffic so accurately in the empirical analysis. This is because the factors that influence air passenger and air cargo demand are different. In addition, the allocation of passenger seat and cargo bay space may not be modified arbitrarily, nor can the fleet.

5. CONCLUSIONS

In the current study, artificial neural networks were employed to forecast air passenger and air cargo demand from Japan to Taiwan. Back-propagation networks can accurately forecast air passenger and air cargo demand, both individually and simultaneously. However, it is more accurate to forecast air passenger and air cargo individually. Seven commonly used criteria were selected to evaluate the forecasting performance of neural networks. *MAPE* is the recommended criterion to evaluate forecasting performance after analyzing the empirical results. The results reveal that some factors influence both air passenger and air cargo volume, while others only influence one of them. In addition, more input variables do not guarantee more accurate forecasts. It is noteworthy that inappropriate input variables will decrease the forecasting accuracy. In addition, a greater number of training years also do not ensure more accurate forecasting. It is thus seen that a sufficient and well-chosen training set is very important in empirical analysis. The procedure used in this study may be adopted as a practical standard operating procedure (SOP) to evaluate input variables.

A contribution graph was also used to study the influence of processing units in the input layer. It is interesting to note that the results obtained by deleting and adding candidate variables are similar with those obtained with a contribution graph. Using the workload unit to reflect air traffic volume is not recommended, as air cargo volume may not be proportional to air passenger volume in empirical analysis. Flight movement from Tokyo (NRT) to Taipei (TPE), PCI in Taiwan, and foreign exchange rate were found to be the three most important factors for air passenger demand forecasting in the empirical analysis. Economic growth rate in Taiwan was found to be the most important factor for air cargo demand forecasting. In addition, it is seen that employed population may be more important than population, although this latter variable was commonly used in the previous research. By using a neural network, novel forecasting models that consider the actual air passenger and air cargo markets were established, and the results show that they can be used an accurate tool to forecast air traffic demand, and the data thus obtained can be a good reference for both air carriers and government authorities.

In the present study, only official annual numerical data were selected to be the candidate variables. Dummy variables were not used, and political effects were not considered. In addition, air fare and oil price were not included in the present model. The results show that forecasting air cargo demand is more difficult than forecasting air passenger demand. This is mainly because the present model only considers a single country-pair. Moreover, the unbalanced air cargo flow in the global supply chain is more complicated than air passenger flow. One possibility for future study could be to include other input variables to improve forecasting accuracy for air cargo demand. The other country-pair, city pair or route may be selected to further validate the BPN model. In addition, a forecasting model using seasonal or monthly data would be valuable if the data are available.

Future research could be conducted to forecast air cargo demand in a global routes network using artificial neural networks. After evaluating the international air cargo network structures, several major air cargo hubs may be selected (Matsumoto, 2007). The output variables may be the air cargo volumes, including import, export, transit and transshipment. The social and economic parameters in each hub, country or region are the candidate input variables. It is supposed that air cargo demand may be more accurately forecasted in a global routes network (Kuo et al., 2010), and that many fruitful results could be obtained if the elaborate model is established.

In addition, Song and Li (2008) reviewed the published studies on tourism demand modelling and forecasting from 2000 to 2007 and suggested that integrating both qualitative and quantitative forecasting approaches may improve forecasting accuracy. Future research thus can try to quantify the qualitative variables in the application of the BPN model. Moreover, the potential effects of crises, disasters and other events on air traffic demand may be further analyzed based on appropriate scenario analysis. It is believed that these efforts may make the present BPN model both more powerful and applicable.

ACKNOWLEDGEMENTS

The authors would like to thank the National Science Council of the R. O. C. (Taiwan) for financially supporting this research under contract no. NSC 98-2221-E-156-006.

REFERENCES

Airbus (2009). http://www.airbus.com/en/corporate/gmf/.

Boeing (2009). http://www.boeing.com/commercial/cmo/.

Brons, M., E. Pels, P. Nijkamp and P. Rietveld (2002). Price elasticities of demand for passenger air travel: a meta-analysis. Journal of Air Transport Management, 8, 165-175.

Central Bank, R. O. C. (Taiwan) (2009). http://www.cbc.gov.tw/.

Cho, V. (2003). A comparison of three different approaches to tourist arrival forecasting. Tourism Management, 24, 323-330.

Directorate General of Budget, Accounting and Statistics of Executive Yuan, R. O. C. (Taiwan) (2009). http://www.dgbas.gov.tw/.

Doganis, R. (1992). The Airport Business. Routledge, London.

Gardiner, J., S. Ison and I. Humphreys (2005). Factors influencing cargo airlines' choice of airport: an international survey. Journal of Air Transport Management, 11, 393-399.

Gately, E. (1996). Neural Networks for Financial Forecasting. Wiley, New York.

Grosche, T., T. Rothlauf and A. Heinzl (2007). Gravity models for airline passenger volume estimation. Journal of Air Transport Management, 13, 175-183.

Haykin, S. (1999). Neural Networks: A Comprehensive Foundation. 2nd ed. Prentice Hall, New Jersey.

Hsu, C. I. and Y. H. Wen (1998). Improved grey prediction models for the trans-Pacific air passenger market. Transportation Planning and Technology, 22, 87-107.

Hsu, C. I. and Y. H. Wen (2000). Application of grey theory and multiobjective programming towards airline network design. European Journal of Operational Research, 127, 44-68.

Jorge-Calderon, J. D. (1997). A demand model for scheduled airline services on international European routes. Journal of Air Transport Management, 3, 23-35.

Kuo, S. Y., L. C. Shiau and Y. P. Chang (2010). Air transport demand forecasting in routes network by artificial neural networks. Journal of Aeronautics Astronautics and Aviation Series B, 42, 67-72. (in Chinese)

Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. Tourism Management, 21, 331-340.

Law, R. and N. Au (1999). A neural network model to forecast Japanese demand for travel to Hong Kong. Tourism Management, 20, 89-97.

Lewis, C. D. (1982). Industrial and Business Forecasting Methods. Butterworths, London.

Lin, C. C. and Y. C. Chen (2003). The integration of Taiwanese and Chinese air networks for direct air cargo services. Transportation Research Part A, 37, 629-647.

Masters, T. (1993). Practical Neural Network Recipes in C++. Academic Press, Boston.

Matsumoto, H. (2004). International urban systems and air passenger and cargo flows: some calculations. Journal of Air Transport Management, 10, 241-249.

Matsumoto, H. (2007). International air network structures and air traffic density of world cities. Transportation Research Part E, 43, 269-282.

Ministry of Economic Affairs of Executive Yuan, R. O. C. (Taiwan) (2009). http://www.moea.gov.tw/.

Ministry of Internal Affairs and Communications, Japan (2009). http://www.soumu.go.jp/.

Ministry of Transportation and Communications of Executive Yuan, R. O. C. (Taiwan) (2009). http://www.motc.gov.tw/.

Ohashi, H., T. S. Kim, T. H. Oum and C. Yu (2005). Choice of air cargo transshipment airport: an application to air cargo traffic to/from Northeast Asia. Journal of Air Transport Management, 11, 149-159.

Song, H. and G. Li (2008), Tourism demand modelling and forecasting—A review of recent research. Tourism Management, 29, 203-220.

Wei, C. H. and Y. C. Yang (1999). A study on transit containers forecast in Kaohsiung port: applying artificial neural networks to evaluating input variables. Journal of the Chinese Institute Transportation, 11, 1-20. (in Chinese)

Zhang, A. (2003). Analysis of an international air-cargo hub: the case of Hong Kong. Journal of Air Transport Management, 9, 123-138.

Zhang, A., Y. V. Hui and L. Leung (2004). Air cargo alliances and competition in passenger markets. Transportation Research Part E, 40, 83-100.

Zhang, A. and Y. Zhang (2002). A model of air cargo liberalization: passenger vs. all-cargo carriers. Transportation Research Part E, 38, 175-191.