Transportation Planning and Technology

ISSN: 0308-1060 (Print) 1029-0354 (Online) Journal homepage: https://www.tandfonline.com/loi/gtpt20

Improved grey prediction models for the trans-pacific air passenger market

Chaug-Ing Hsu & Yuh-Horng Wen

Transportation Planning and Technol., Vol. 22, pp. 87–107
Reprints available directly from the publisher
Photocopying permitted by license only

IMPROVED GREY PREDICTION MODELS FOR THE TRANS-PACIFIC AIR PASSENGER MARKET

CHAUG-ING HSU* and YUH-HORNG WEN

Department of Transportation Engineering and Management, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu, Taiwan, 30050, R.O.C.

(Received 7 April 1997; In final form 2 March 1998)

The rapid economic growth of Asia-Pacific countries continues to result in faster travel growth in the trans-Pacific air passenger market. Grey theory is used to develop time series GM(1,1) models for forecasting total passenger and 10 country-pair passenger traffic flows in this market. The accumulated generating operation (AGO) is one of the most important characteristics of grey theory, and its main purpose is to reduce the randomness of data. The original GM(1,1) models are improved by using residual modifications with Markov-chain sign estimations. These models are shown to be more reliable by posterior checks and to yield more accurate prediction results than ARIMA and multiple regression models. The results indicate that the total number of air passengers in the trans-Pacific market will increase at an average annual growth rate of approximately 11% up to the year 2000.

Keywords: Grey theory; Grey model (GM); Air passenger traffic; Accumulated generating operation (AGO); Prediction

1. INTRODUCTION

The rapid economic growth in Asia-Pacific countries is stimulating international travel and will in all probability result in even faster international travel growth in the future. The trans-Pacific passenger market is defined by the Boeing Commercial Airplane Group as including all passengers traveling on scheduled and charter flights

* Corresponding author. Fax: 886-3-5720844.

between the United States (U.S.) and countries including Hong Kong, Indonesia, Japan, South Korea (Korea), Malaysia, People's Republic of China (China), Philippines, Singapore, Republic of China (Taiwan) and Thailand (Boeing, 1993). The trans-Pacific passenger market is one of the fastest growing air travel markets in the world. This paper attempts to apply grey theory to develop improved grey prediction models for this market.

The need for accurate passenger forecasts in the commercial airline industry is extensive due to its importance and numerous practical applications. Airlines use short-term forecasts for daily operations, medium-term forecasts for market planning, and long-term forecasts for route and fleet planning. Airports use passenger throughput forecasts for terminal area and access facility planning. Government agencies use such forecasts for aviation system budgeting and policy formulation (Huth and Eriksen, 1987; Horonjeff and McKelvey, 1994).

The methodologies for predicting air traffic presented in the literature are as diverse as the applications themselves. Frequently used quantitative techniques include time-series and causal methods. Causal methods include regression and econometric techniques where relationships between air traffic and a set of causal explanatory variables are estimated.

Casual methods (e.g., FAA, 1980; Ippolito, 1981; Kaemmerle, 1991; Russon and Riley, 1993; Boeing, 1993; 1995) and time-series methods (e.g., Young, 1972; Behbehani and Kanafani, 1980; Huth and Eriksen, 1987) have both been used to investigate and forecast city-pair air passenger and country total air passenger traffic patterns. However, the well-known assumptions underlying these two types of methods limit their application validity. Other methods, including the application of neural networks to forecasting international airline passenger traffic have been developed (Nam and Schaefer, 1995). Neural networks do not require making as many assumptions as these other statistical methods, but they still require a large number of consistent data sets for training and to establish a learning procedure.

Economic conditions from 1970 to the present have varied widely from region to region. The relative growth of traffic in the Asian and Pacific region is quite dramatic, paralleling the growth of importance of this area in the political, social and economic sectors during the same period (Boeing, 1995; Horonjeff and McKelvey, 1994). However, the number of yearly observations has been small, because the fast growth of the trans-Pacific air passenger market has occurred only over recent decades. As a matter of fact, the numbers of observations are usually small, especially for country-pair data. Therefore, it is difficult to collect large quantities of data points with good enough statistical distributions to construct satisfactory conventional statistical models. We thus attempt to apply grey theory to develop forecasting models for international air travel market.

Grey theory was developed originally by Deng (1982). Although much of the literature is in Chinese (owing to its origin), its principles and application areas are presented and discussed in English by Deng (1988; 1989). It is a truly multidisciplinary and generic theory which deals with systems which are characterised by poor information and/or for which information is lacking. The fields covered by grey theory include systems analysis, data processing, modelling, prediction, decision making and control. Its applications are numerous, as any issue of the *Journal of Grey System* will testify. In contrast to traditional statistical methods, the potency of the original series in the time-series grey model, called GM(1,1), has been proven to be more than four (Deng *et al.*, 1988). In addition, assumptions regarding statistical distribution of data are not necessary when applying grey theory. The accumulated generating operation (AGO) is one of the most important characteristics of grey theory, and its main purpose is to reduce the randomness of data. In fact, functions derived from AGO formulations of original series are always well-fitted to exponential functions (Deng, 1982). Grey prediction models have been recently used in many applications (e.g., Deng *et al.*, 1988; Sun, 1991; Chen and Tien, 1993).

This paper is the first attempt to apply grey theory to international air passenger prediction modeling. We develop time series GM(1,1) models for forecasting total passenger traffic and 10 country-pair passenger traffic in the trans-Pacific market. And we introduce a new technique that combines residual modification and residual Markov-chain sign estimation to improve the accuracy of the original models. The results of these improved models are tested to examine reliability; and they are compared with those of multiple regression models and autoregressive integrated moving-average (ARIMA) models.

2. GREY MODEL FOR TIME-SERIES FORECASTING

The grey forecasting model GM(1,1) is a time-series prediction model encompassing a group of differential equations adapted for parameter variance than a first-order differential equation. Its difference equations have structures that vary with time rather than being general difference equations. Although it is not necessary to employ all the data from the original series to construct GM(1,1), the potency of the series must be more than 4. In addition, the data must be taken at equal intervals and in consecutive order without bypassing any data (Deng *et al.*, 1988).

Assume an original total annual passenger series to be

$$Y^{(0)} = \left[Y^{(0)}(1), Y^{(0)}(2), \ldots, Y^{(0)}(n)\right], \tag{1}$$

where n stands for the number of years observed. The AGO formation of $Y^{(0)}$ will be

$$Y^{(1)} = \left[Y^{(1)}(1), Y^{(1)}(2), \ldots, Y^{(1)}(n)\right], \tag{2}$$

where $Y^{(1)}(1) = Y^{(0)}(1)$,

$$Y^{(1)}(k) = \sum_{i=1}^{k} Y^{(0)}(i), \quad k = 2, 3, \ldots, n. \tag{3}$$

The GM(1,1) model can be constructed by establishing a first-order differential equation for $Y^{(1)}$. That is

$$\frac{\mathrm{d}Y^{(1)}(k)}{\mathrm{d}k} + aY^{(1)}(k) = u. \tag{4}$$

The solution of Eq. (4) can be obtained using the least-squares method. That is

$$\hat{Y}^{(1)}(k+1) = \left(Y^{(0)}(1) - \frac{u}{a}\right)\mathrm{e}^{-ak} + \frac{u}{a}, \tag{5}$$

where

$$[a, u]^{\mathrm{T}} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}y_N \tag{6}$$

and

$$B = \begin{bmatrix} -\frac{1}{2}(Y^{(1)}(1) + Y^{(1)}(2)), & 1 \\ -\frac{1}{2}(Y^{(1)}(2) + Y^{(1)}(3)), & 1 \\ \vdots & \vdots \\ -\frac{1}{2}(Y^{(1)}(n-1) + Y^{(1)}(n)), & 1 \end{bmatrix}, \tag{7}$$

$$y_N = \left[Y^{(0)}(2), Y^{(0)}(3), \ldots, Y^{(0)}(n)\right]^{\mathrm{T}}. \tag{8}$$

We obtain $\hat{Y}^{(1)}$ from Eq. (5). Let $\hat{Y}^{(0)}$ be the fitted and forecast series,

$$\hat{Y}^{(0)} = \left[\hat{Y}^{(0)}(1), \hat{Y}^{(0)}(2), \ldots, \hat{Y}^{(0)}(n), \ldots\right] \tag{9}$$

and $\hat{Y}^{(0)}(1) = Y^{(0)}(1)$. Applying the inverse accumulated generating function (IAGO), we then have

$$\hat{Y}^{(0)}(k) = \left(Y^{(0)}(1) - \frac{u}{a}\right)(1 - \mathrm{e}^{a})\mathrm{e}^{-a(k-1)}, \quad k = 2, 3, \ldots, \tag{10}$$

where $\hat{Y}^{(0)}(1)$, $\hat{Y}^{(0)}(2)$, $\ldots$, $\hat{Y}^{(0)}(n)$ are the fitted values of the original series and $\hat{Y}^{(0)}(n+1)$, $\hat{Y}^{(0)}(n+2), \ldots,$ are the forecast values.

In this paper, we use annual passenger totals for the trans-Pacific air market from 1974 ($k=1$) through 1993 ($k=20$) as reported by Boeing (1993) for our data. There are 20 observations, of which the first 18 are used for model fitting and the last two are reserved for post-sample result comparisons. We construct matrixes B and y_N in Eqs. (7) and (8) and apply Eq. (6) to give $[a,u]^{\mathrm{T}} = [-0.10552, 1895.225]^{\mathrm{T}}$. Then from Eq. (10), there follows

$$\hat{Y}^{(0)}(k) = \left(2026.97 + \frac{1895.225}{0.10552}\right)(1 - \mathrm{e}^{-0.10552})\mathrm{e}^{0.10552(k-1)}, \quad k = 2, 3, \ldots, \tag{11}$$

where $\hat{Y}^{(0)}(2)$, $\hat{Y}^{(0)}(3)$, $\hat{Y}^{(0)}(4)$, $\ldots$, are forecast values for 1974 ($k=2$), 1975 ($k=3$), 1976 ($k=4$), $\ldots$, and $\hat{Y}^{(0)}(1)$ is the original value for 1974 ($Y^{(0)}(1)$). The data years, real values, model values, and forecast percentage errors are shown in the first four columns of Table I.

TABLE I The model values and forecast errors of both original model and improved model

Year	*Real value*	*Original model*		*Limits*		*Improved model*	
		Model value	*Error* (%)	*Lower limits*	*Upper limits*	*Model value*	*Error* (%)
1974	2026.97	2026.97	—	1800	2300	2026.97	—
1975	2155.1	2224.409	−3.216	2023.75	2465.16	2104.835	2.33
1976	2571.85	2471.961	3.884	2243.52	2739.29	2602.223	−1.181
1977	2652.84	2747.064	−3.551	2487.16	3043.91	2605.16	1.797
1978	2997.69	3052.782	−1.837	2757.25	3382.41	2898.196	3.32
1979	3726.28	3392.523	8.957	3056.67	3758.54	3560.926	4.437
1980	4164.16	3770.074	9.464	3388.61	4176.51	3953.529	5.058
1981	4419.58	4189.643	5.203	3756.59	4640.95	4389.494	0.681
1982	4624.72	4655.904	−0.67	4164.54	5157.04	4438.191	4.033
1983	4778.29	5174.056	−8.283	4616.78	5730.52	4936.884	−3.319
1984	5662.37	5749.872	−1.545	5118.14	6367.77	5491.502	3.018
1985	6096.65	6389.77	−4.81	5673.94	7075.89	6108.308	−0.191
1986	6808.1	7100.882	−4.3	6290.1	7862.75	6794.264	0.203
1987	7795.7	7891.132	−1.22	6973.18	8737.12	7557.11	3.061
1988	9068	8769.329	3	7730.42	9708.72	9133.206	−0.719
1989	10194.7	9745.26	4.41	8569.91	10788.4	10141.66	0.52
1990	11394.6	10829.8	4.956	9500.55	11988.1	11261.63	1.167
1991	11588	12035.04	−3.858	10532.3	13321.2	11564.62	0.2
1992	12844	13374.41	−4.13	11676	14802.5	12861.94	−0.14
1993	13610.3	14862.84	−9.203	12944	16448.6	14304.56	−5.1

The upper and lower limits of the original and forecast series capture the extent of the variation in annual air passenger traffic evolution trends. The upper and lower series limits can be estimated separately by applying GM(1,1) modeling, that is, Eqs. (1)–(10), using, respectively, the upper and lower points and other appropriately selected points along the boundaries of the original series. These limits are shown in the fifth and sixth columns of Table I and also plotted in Fig. 1.

3. IMPROVED GREY MODEL USING RESIDUAL MODIFICATION AND MARKOV-CHAIN SIGN ESTIMATION

The differences between real values, $Y^{(0)}(k)$, and model-fitted values, $\hat{Y}^{(0)}(k)$, are called the residual series. The GM(1,1) of a residual series may be established to further improve predictive accuracy (Deng, 1982).

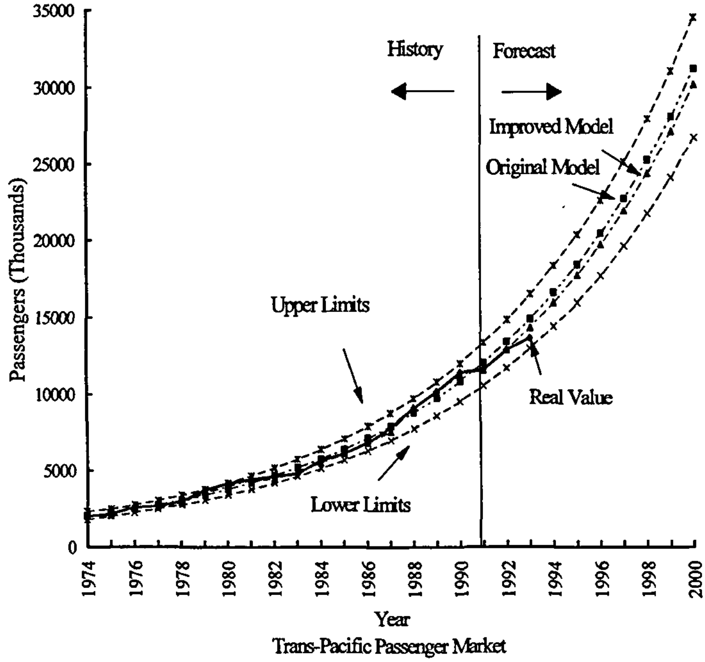

FIGURE 1 The real values, upper and lower limits, and model values for total passenger traffic from 1974 to 2000.

The new predicted values can be obtained by adding the forecast values of residual GM(1,1) to $\hat{Y}^{(0)}(k)$ in Eq. (11) for each year. However, the potency of the residual series depends on the number of data points with the same sign, which is usually small when there are few observations. In these cases, the potency of the residual series with the same sign may not be more than 4, so a residual GM(1,1) cannot be established.

Here, we present an improved grey model to solve this problem. We establish a GM(1,1) using the absolute values of the residual series such that the potency of the series is more than 4. Then we set up a two-state Markov chain model to predict the sign of the forecast residual, where state 1 denotes a positive sign and state 2 denotes a negative sign.

Denote the residual series, the difference between the real $Y^{(0)}(k)$ and the model-fitted $\hat{Y}^{(0)}(k)$ obtained from Eq. (11), as $q^{(0)}$:

$$q^{(0)} = \left[q^{(0)}(2), q^{(0)}(3), \ldots, q^{(0)}(n)\right], \tag{12}$$

where

$$q^{(0)}(k) = Y^{(0)}(k) - \hat{Y}^{(0)}(k). \tag{13}$$

Denote the absolute values of the residual series as $\varepsilon^{(0)}$:

$$\varepsilon^{(0)} = \left[\varepsilon^{(0)}(2), \varepsilon^{(0)}(3), \ldots, \varepsilon^{(0)}(n)\right], \tag{14}$$

where

$$\varepsilon^{(0)}(k) = \left|q^{(0)}(k)\right|, \quad k = 2, 3, \ldots, n. \tag{15}$$

A GM(1,1) model of $\varepsilon^{(0)}$ can be established by using methods similar to those described in Section 2, then

$$\hat{\varepsilon}^{(0)}(k) = (0.08561)\left(119.5741 + \frac{113.6408}{0.08561}\right)e^{0.08561(k-1)}, \quad k = 2, 3, \ldots, n, \ldots \tag{16}$$

A Markov-chain model is used here to predict the signs of the forecast residual series; e.g., $\hat{\varepsilon}^{(0)}(k)$, $k = n+1, n+2, \ldots$. Assume if the sign of the kth year residual is positive, it is in state 1, and that if it is in state 2, the sign is negative. With each possible transition from state i to state j, we associate a one-step transition probability, P_{ij}. P_{ij} can be estimated using

$$P_{ij} = \frac{M_{ij}}{M_i}, \quad i = 1, 2 \text{ and } j = 1, 2, \tag{17}$$

where M_i is the number of years whose residuals are in state i, and M_{ij} is the number of transitions from state i to state j that have occurred. These P_{ij} values can be arranged as a transition matrix, R,

$$R = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}. \tag{18}$$

R can be estimated by examining the signs of residuals for all years from 1975 to 1991. Of the 7 years with positive residuals previously, 4 (or 57.14%) stayed positive, whereas 3 (or 42.86%) changed to negative. Similarly, of those with a negative sign, a third changed to positive, and two-thirds stayed negative. The estimated transition matrix, $\hat{R}$, associated with this process is

$$\hat{R} = \begin{bmatrix} 0.571 & 0.429 \\ 0.333 & 0.667 \end{bmatrix}. \tag{19}$$

$\hat{R}$ can be used to predict the probability of a residual sign being in state 1 or state 2 in future years. Denote the initial state distribution by the vector $\pi^{(0)} = [\pi_1^{(0)}, \pi_2^{(0)}]$, and assume year 1991 is the initial state. We then have $\pi^{(0)} = [0, 1]$ representing its residual being negative. With $\pi^{(0)}$ as the initial state distribution, the state probabilities after n' transitions are given by

$$\pi^{(n')} = \pi^{(0)} \hat{R}^{n'}, \quad n' = 1, 2, \ldots, \tag{20}$$

where $\pi^{(n')} = [\pi_1^{(n')}, \pi_2^{(n')}]$. Thus, the state probabilities for 1991 and following years can be estimated using Eq. (20). Let the sign of the kth year residual, $\delta(k)$, be

$$\delta(k) = \begin{cases} +1, & \text{if it is positive.} \\ -1, & \text{if it is negative.} \quad k = 1, 2, \ldots, n, \ldots \end{cases} \tag{21}$$

Then, if $\pi_1^{(n')} > \pi_2^{(n')}$, the residual sign of the year is positive; e.g., $\delta(n+n') = +1$; otherwise, it is negative, e.g., $\delta(n+n') = -1$. From Eqs. (11), (16), (19), (20) and (21), an improved grey model with residual modification and Markov-chain sign estimation can be formulated. That is,

$$\begin{aligned} \hat{Y}'^{(0)}(k) &= \left(2026.97 + \frac{1895.225}{0.10552}\right)(1 - e^{-0.10552})e^{0.10552(k-1)} \\ &\quad + \delta(k)(0.08561)\left(119.5741 + \frac{113.6408}{0.08561}\right)e^{0.08561(k-1)}, \\ &\quad k = 1, 2, \ldots, n, n+1, \ldots \end{aligned} \tag{22}$$

The results of applying this improved grey model are shown in columns 7 and 8 of Table I. Table I also shows the forecast percentage errors of the improved grey model to be lower than those of the original grey model.

The posterior check is used to test the reliability of the forecast model. Let S_1 and S_2 be, respectively, the standard deviations of original series $Y^{(0)}(k)$, and residual series $q^{(0)}(k)$. According to Deng (1986), the posterior check ratio, C, is defined as

$$C = \frac{S_2}{S_1}, \tag{23}$$

and the frequency of the small error, p, is

$$p = p\left\{ \left| q^{(0)}(k) - \overline{q} \right| < 0.6745 S_1 \right\}, \tag{24}$$

where $\overline{q}$ is the average value of the residual series $q^{(0)}(k)$. The judge of precision accuracy grade, JPA, is based on overall evaluations of p and C values according to the criteria listed in Table II.

The grade of precision accuracy is higher when the C value is smaller, because a smaller value implies a larger S_1 value and a smaller S_2 value. That is, though the original series may have large variations and poor regularity, the residual series of the forecast model have a small standard deviation.

On the other hand, the higher the p value, the higher the precision accuracy grade is. Higher values indicate larger proportions of data points with residual values less than the average residual by values smaller than 0.6745.

The results of posterior checks of the improved model and the original model are shown in Table III. The C value of the improved model is lower than that of the original model, though both models are evaluated as "good". This indicates that the improved grey model is slightly more reliable than the original grey model.

TABLE II Judging for the grade of precision accuracy, JPA

JPA	p	C
Grade 1: Good	$p \geqq 0.95$	$C \leqq 0.35$
Grade 2: Qualified	$0.95 \geqq p \geqq 0.8$	$0.35 \leqq C \leqq 0.5$
Grade 3: Just the mark	$08 \geqq p \geqq 0.7$	$0.5 \leqq C \leqq 0.65$
Grade 4: Unqualified	$p < 0.7$	$C > 0.65$

TABLE III The JPA of original model and improved models

Original model	*Improved model*
$C=0.097172<0.35$ (Good)	$C=0.033919<0.35$ (Good)
$p=1>0.95$ (Good)	$p=1>0.95$ (Good)

The forecast errors of the improved grey model are also compared with those of two other statistical models. One is a multiple regression model described in a series of reports published by Boeing (1993; 1995), and the other is an ARIMA model (Box *et al.*, 1994). The multiple regression model used by Boeing (1993) is

$$\log(Y) = -12.04 + 1.23\log(x_1) + 1.51\log(x_2) - 0.11\log(x_3), \quad (25)$$

where x_1 is Gross Domestic Product (GDP) of U.S., in billions of 1980 dollars, x_2 is GDP of Asia-Pacific countries, in billions of 1980 dollars, and x_3 is average yields in 1980 prices, in cents per revenue passenger mile (RPM).

We use the same number of observations, 18 (from 1974 to 1991), to formulate an ARIMA(p, d, q) model, where p is the order of the auto-regressive part, d is the order of the differencing, and q is the order of the moving-average process (Box *et al.*, 1994). As a result of statistical tests, an ARIMA model with $(p, d, q) = (0, 1, 0)$ is formulated as follows:

$$\log(\hat{Y}(k)) = 0.1078775 + \log(\hat{Y}(k-1)), \quad k = 2, 3, \ldots, n, \ldots \quad (26)$$

and $\hat{Y}(1) = Y(1)$.

The results obtained by the improved grey model, ARIMA model and Boeing multiple regression model are shown in Table IV. These results as well as forecasts for years from 1994 to 2000 are also plotted in Fig. 2. The improved grey model is shown to yield the lowest forecasting errors among the three models. Thus, the grey model is a powerful forecasting model, especially when the number of observations is not large as is the case for the trans-Pacific air passenger market. It is convenient to use the original grey model GM(1,1) to estimate annual growth rate, because $\hat{Y}^{(0)}(k+1)/\hat{Y}^{(0)}(k) = e^{0.10552} = 1.1113$ is a constant. An annual growth rate of 11.13% is obtained by this method. On the other hand, an average rate of 11.06% is obtained by calculating the

TABLE IV The model values and forecast errors for the improved model, Boeing model, and ARIMA model

Year	Real value	Improved model		Boeing model		ARIMA model	
		Model value	Error (%)	Model value	Error (%)	Model value	Error (%)
1974	2026.97	2026.97	—	2061.58	−1.71	2026.97	—
1975	2155.1	2104.835	2.33	2117.558	1.742	2257.865	−4.768
1976	2571.85	2602.223	−1.181	2427.397	5.617	2515.061	2.208
1977	2652.84	2605.16	1.797	2803.448	−5.68	2801.555	−5.606
1978	2997.69	2898.196	3.32	3244.388	−8.23	3120.684	−4.103
1979	3726.28	3560.926	4.437	3665.776	1.624	3476.165	6.712
1980	4164.16	3953.529	5.058	3922.363	5.807	3872.139	7.013
1981	4419.58	4389.494	0.681	4292.079	2.885	4313.219	2.407
1982	4624.72	4438.191	4.033	4399.542	4.869	4804.544	−3.888
1983	4778.29	4936.884	−3.319	4854.996	−1.61	5351.835	−12.003
1984	5662.37	5491.502	3.018	5745.64	−1.47	5961.47	−5.282
1985	6096.65	6108.308	−0.191	6424.846	−5.383	6640.549	−8.921
1986	6808.1	6794.264	0.203	6975.245	−2.455	7396.982	−8.65
1987	7795.7	7557.11	3.061	7863.791	−0.873	8239.582	−5.694
1988	9068	9133.206	−0.719	9056.336	0.129	9178.163	−1.215
1989	10194.7	10141.66	0.52	10087.4	1.052	10223.66	−0.284
1990	11394.6	11261.63	1.167	11161.26	2.048	11388.25	0.056
1991	11588	11564.62	0.2	11770.69	−1.577	12685.5	−9.471
1992	12844	12861.94	−0.14	13510.06	−5.186	14130.52	−10.02
1993	13610.3	14304.56	−5.1	14631.39	−7.431	15740.15	−15.57

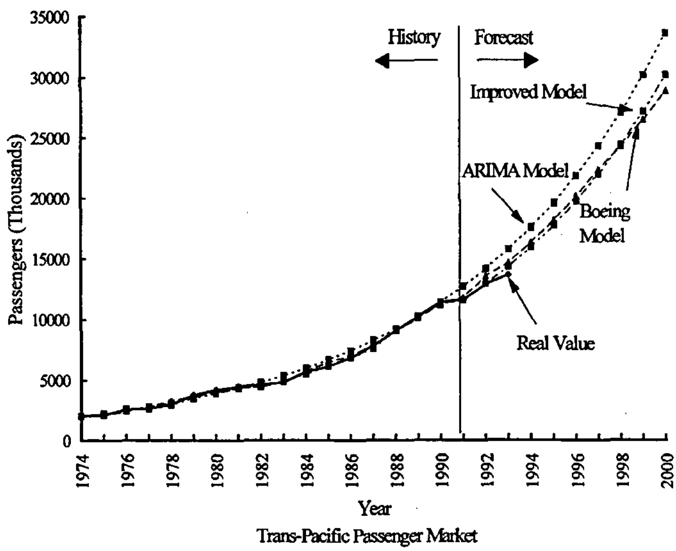

FIGURE 2 The real values, model values for the improved model, ARIMA model and Boeing model for total passenger traffic from 1974 to 2000.

average of annual growth rates using the forecast values of the improved grey model. These results indicate that the total number of air passengers in the trans-Pacific market is likely to grow at an annual rate of roughly 11% in the coming years.

4. FORECASTING TRANS-PACIFIC AIR PASSENGER TRAFFIC BY COUNTRY

In addition to the total passenger forecast, this section presents trans-Pacific country-pair passenger forecasts using modeling techniques similar to those described in Sections 2 and 3. We formulated GM(1,1) models, estimated upper and lower limits, and set up improved GM(1,1) models with residual modifications and Markov-chain sign estimations for all country-pair passenger forecasts. All models were then tested to determine whether they were reliable using posterior checks. Furthermore, the Boeing forecasts (Boeing, 1995) were used to compare the predictive accuracy of these models.

In this process, we used historic trans-Pacific air passenger data listed in reports by Boeing (1993). The number of available data points for historic passenger traffic between the U.S. and each of 10 Asia-Pacific countries was far lower than that for total trans-Pacific passenger traffic. Nine observations were available for the years 1983 to 1991, for model fitting regarding 10 countries, excluding Malaysia and Indonesia, for which only 6 observations were available for the years 1986 to 1991. Observations for the years 1992 and 1993 were reserved for post-sample comparisons.

Table V shows the estimated coefficients, $\hat{a}$ and $\hat{u}$, for the original and residual GM(1,1) for forecasting trans-Pacific air passenger traffic by country. The results of posterior checks for these models are presented in Table VI. From the original grey models for 10 countries, four (those for Japan, Hong Kong, Korea and Thailand) were considered "good", while five (those for Taiwan, Philippines, Singapore, Malaysia and Indonesia) were considered "qualified", and one, China, was considered "just the mark" or "unqualified". By contrast, all of the improved grey models for the 10 countries were considered "good". The improved grey models with residual modifications and Markov-chain sign estimations appear to have significantly higher grades of precision accuracy than

TABLE V Estimated coefficients, $\hat{a}$ and $\hat{u}$, for the original and residual GM(1,1) in forecasting trans-Pacific air passenger traffic by country

Country	*Original GM (1, 1) model*		*Residual GM (1, 1) model*	
	$\hat{a}$	$\hat{u}$	$\hat{a}$	$\hat{u}$
Japan	−0.11127	3216.059	−0.15133	199.4043
Hong Kong	−0.05211	474.7024	−0.14584	2.092718
Korea	−0.13529	510.1847	−0.06567	49.01213
Taiwan	−0.12188	264.4206	−0.10074	37.1684
Philippines	−0.06937	237.7649	0.07604	30.8248
Singapore	−0.14099	60.97438	−0.28813	1.489082
China	−0.02909	67.04808	−0.26547	2.407199
Thailand	−0.15155	25.34779	−0.17343	2.043745
Malaysia	−0.25889	18.73714	−0.25308	5.286708
Indonesia	−0.18445	29.4253	−0.27169	0.546389

TABLE VI The results of posterior checks for the original model and improved model by country

Country	*Original model*		*Improved model*	
	C	p	C	p
Japan	0.228514 < 0.35 (Good)	1 > 0.95 (Good)	0.073982 < 0.35 (Good)	1 > 0.95 (Good)
Hong Kong	0.264117 < 0.35 (Good)	1 > 0.95 (Good)	0.144087 < 0.35 (Good)	1 > 0.95 (Good)
Korea	0.238755 < 0.35 (Good)	1 > 0.95 (Good)	0.124706 < 0.35 (Good)	1 > 0.95 (Good)
Taiwan	0.394329 < 0.5 (Qualified)	0.875 > 0.8 (Qualified)	0.14824 < 0.35 (Good)	1 > 0.95 (Good)
Philippines	0.411985 < 0.5 (Qualified)	0.875 > 0.8 (Qualified)	0.177827 < 0.35 (Good)	1 > 0.95 (Good)
Singapore	0.462512 < 0.5 (Qualified)	0.875 > 0.8 (Qualified)	0.151725 < 0.35 (Good)	1 > 0.95 (Good)
China	0.64149 < 0.65 (Just the mark)	0.5 < 0.7 (Unqualified)	0.14127 < 0.35 (Good)	1 > 0.95 (Good)
Thailand	0.243409 < 0.35 (Good)	1 > 0.95 (Good)	0.080387 < 0.35 (Good)	1 > 0.95 (Good)
Malaysia	0.496676 < 0.5 (Qualified)	0.8 ≥ 0.8 (Qualified)	0.064986 < 0.35 (Good)	1 > 0.95 (Good)
Indonesia	0.352961 < 0.5 (Qualified)	1 > 0.95 (Good)	0.131508 < 0.35 (Good)	1 > 0.95 (Good)

most of the original grey models. That is, though the original series have large variations and poor regularity, the residual series from the improved model yield small standard deviations and higher proportions of data points with residuals less than the average by values smaller than 0.6745.

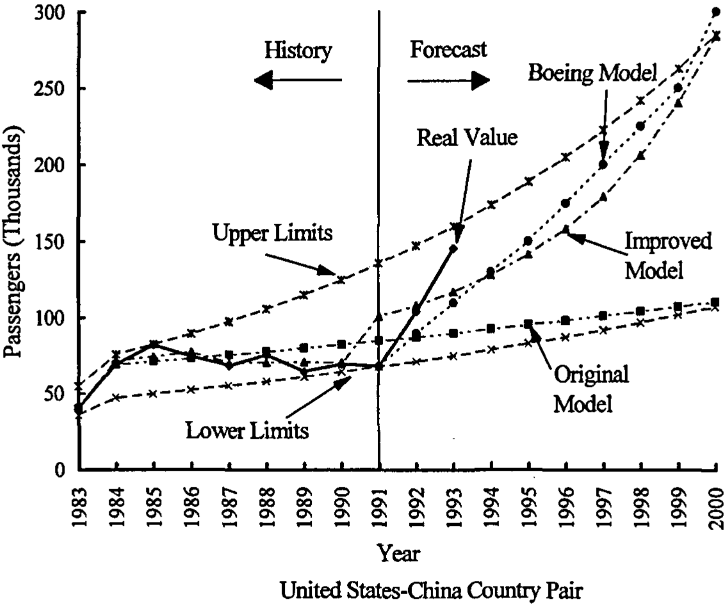

FIGURE 3 The real values, upper and lower limits, and model values for the improved model and Boeing model for U.S.–China country-pair traffic from 1974 to 2000.

Of 10 Asia-Pacific countries, China's improved grey model shows the most significant enhancement. Air passenger traffic between the U.S. and China experienced considerable variation resulting from fluctuations in political conditions, such as those associated with the Tienanmen Square events of 1989, as shown in Fig. 3. On the other hand, rapid economic growth in China fueled increased traffic beginning in 1992 (see Fig. 3) as a result of policies aimed at accelerating economic reformation and opening domestic markets inaugurated in 1992. As shown in Table VI, the original grey model for China has a large C $(0.65 > 0.5)$ value and a small p $(0.5 < 0.7)$ value since considerable variation in historic data resulted in half the model values having unacceptable residuals. Nevertheless, the improved grey model indeed had greatly enhanced prediction accuracy over the original model.

Boeing (1993) used the "share forecast" technique to derive passenger forecasts between the U.S. and each of 10 Asia-Pacific countries. Table VII compares the forecasting error percentages between the results from the improved grey model and those of the Boeing model.

TABLE VII Model values and forecasting errors for the original model, improved model and Boeing model by country

Country	*Year*	*Real value*	*Original model*		*Improved model*		*Boeing model*	
			Model value	*Error* (%)	*Model value*	*Error* (%)	*Model value*	*Error* (%)
Japan	1992	8313.4	9202.874	−10.7	8517.717	−2.46	9060	−8.98
	1993	8488	10286.03	−21.18	9488.929	−11.79	9830	−15.81
Hong Kong	1992	879.2	771.8552	12.21	799.553	9.05	787	10.49
	1993	961.5	813.146	15.43	845.192	12.1	826	14.09
Korea	1992	1729.8	1808.477	−4.55	1715.86	0.81	1640	5.19
	1993	1874.7	2070.479	−10.44	1971.57	−5.17	1777	5.21
Taiwan	1992	888.1	841.522	5.24	892.651	−0.51	888	0.011
	1993	1160	950.6	18.05	1047.398	9.71	958	17.41
Philippines	1992	457.8	459.6506	−0.4	459.651	−0.4	449	1.92
	1993	457.4	492.6684	−7.71	492.668	−7.71	476	−4.07
Singapore	1992	195.6	254.913	−30.32	210.9331	−7.84	217	−10.94
	1993	272.6	293.51	−7.67	268.3676	1.55	235	13.79
China	1992	104	87.3698	15.99	108.1037	−3.95	90	13.46
	1993	145.5	89.949	38.18	116.9869	19.6	110	24.4
Thailand	1992	144.8	110.924	23.4	120.498	16.78	103	28.87
	1993	110	129.075	−17.34	117.687	−6.99	113	−2.72
Malaysia	1992	67.8	87.3102	−28.78	65.945	2.74	62	8.55
	1993	85	113.11	−33.07	85.5913	−0.7	70	17.65
Indonesia	1992	63.6	82.13328	−29.14	68.32	−7.42	101	−58.81
	1993	64.6	98.76987	−52.89	80.644	−24.84	117	−81.11

This comparison used real values from two observations for the years 1992 and 1993 as a basis. The percentages of forecasting errors for the improved grey model apparently were lower than those of forecasts made by the Boeing model for most countries.

Figures 3–6 illustrate, respectively, the actual data, forecasts from the improved grey model, and the Boeing model, and the upper and lower limits for air passenger traffic between the U.S. and China, Japan, Hong Kong and Taiwan. These figures depict the past and future trends in passenger traffic between the U.S. and each of four Asia-Pacific countries. The variations in trends for different countries depend on the individual countries' economic conditions.

Table VIII shows the average growth rates for historic data and for future forecasts by country. Higher growth-rate countries are shown to include Korea, Taiwan, Thailand, Malaysia and Indonesia according to

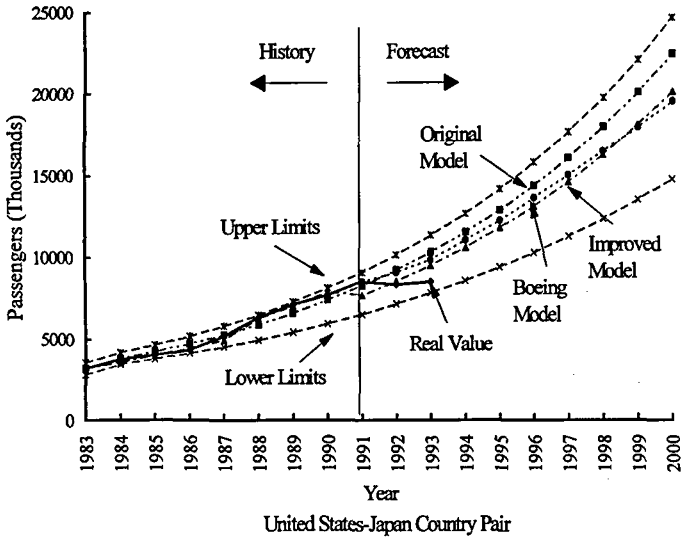

FIGURE 4 The real values, upper and lower limits, and model values for the improved model and Boeing model for U.S.–Japan country-pair traffic from 1974 to 2000.

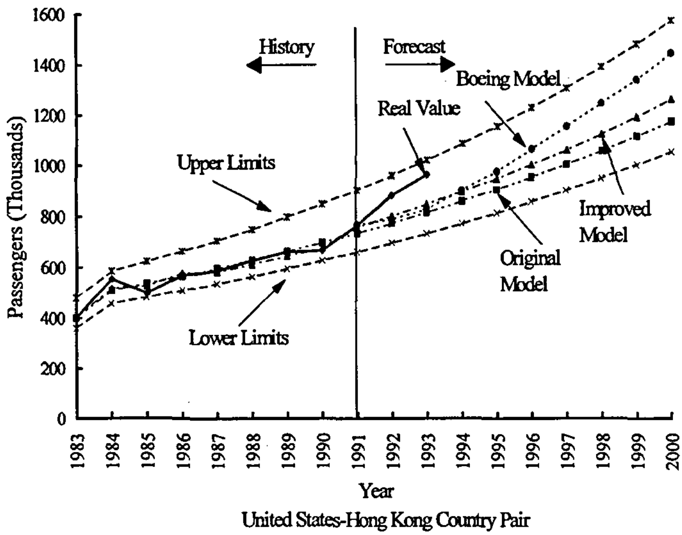

FIGURE 5 The real values, upper and lower limits, and model values for the improved model and Boeing model for U.S.–Hong Kong country-pair traffic from 1974 to 2000.

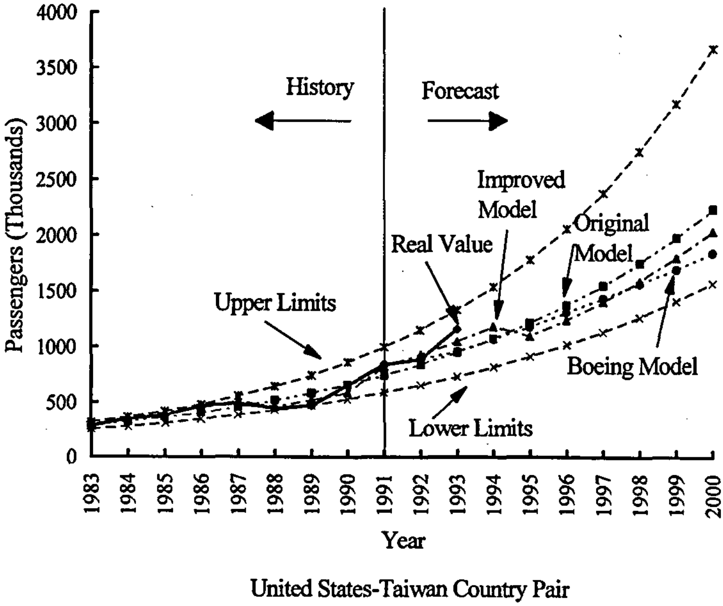

FIGURE 6 The real values, upper and lower limits, and model values for the improved model and Boeing model for U.S.–Taiwan country-pair traffic from 1974 to 2000.

history. They all have annual growth rates larger than 12%. Passenger traffic from Korea and Taiwan to the U.S. has increased substantially. In addition to economic growth, the entry of new airlines such as Korea's *Asiana Airline* and Taiwan's *EVA Airline* also led to significantly higher outbound traffic from these two countries. On the other hand, Thailand, Malaysia and Indonesia are experiencing the highest economic growth, though they started with low levels of passenger traffic. Nevertheless, the average annual growth rate for the years 1992–2000 is predicted to gradually decline for these countries. China has experienced rapid economic growth since 1991, thus it is expected to have an annual growth rate of roughly 12% in the years 1992–2000.

By comparison, the share-forecast technique used by Boeing assumes predicted total passenger traffic on the basis of the *a priori* multiple regression model in Eq. (25), so forecast errors tend to be larger when the multiple regression model does not produce sufficient precise results. On the other hand, forecasts using the time-series technique demand large amounts of data. Conversely, it appears that the

TABLE VIII Average annual growth rate by country

Country	*Average growth rate* (%)	
	Historic data	*Forecast* (1992–2000)
Japan	11.66	11.4
Hong Kong	9.19	5.69
Korea	17.01	14.62
Taiwan	15.6	10.58
Philippines	8.05	7.1
Singapore	10.65	12.84
China	9.77	11.87
Thailand	15.38	14.6
Malaysia	35.26	28.36
Indonesia	33.33	16.54

improved grey models presented in this paper do not have these disadvantages, but still produce more accurate results. However, the state probabilities in Markov-chain prediction tend to converge when the number of transitions, i.e., number of forecast years, are large. Therefore, the improved grey model with residual modification and Markov-chain sign estimation is not suitable for forecasting future traffic over very long periods of time. The modeling process of the improved GM(1,1) model is simple. Forecasters can regularly update improved GM(1,1) models by entering new data to improve forecasting accuracy to designed levels.

5. CONCLUSIONS

This paper has applied grey theory to develop time series GM(1,1) models for predicting traffic flows in the trans-Pacific air passenger market. The GM(1,1) model is a dynamic model with a group of differential equations adapted for variance of parameters. The original GM(1,1) models were further improved by using a technique that combines residual modifications with Markov-chain sign estimations. This technique not only caused the improved model to yield more accurate results but also solved problems resulting from having too few samples, which may make the potency of the same sign residuals lower than four and violate the necessary condition of the GM(1,1) model. The improved grey models were then applied to forecast trans-Pacific total passenger traffic and 10 country-pair passenger traffic flows.

According to posterior checks, all of the improved grey models were considered "good" and reliable. In addition, the improved grey models were also shown to yield more precise forecasts than conventional statistical models such as ARIMA and multiple regression.

Our results indicate that the total number of air passengers in the trans-Pacific market will probably grow at an average annual growth rate of approximately 11% up to the year 2000. Historic trends indicate Korea, Taiwan, Thailand, Malaysia and Indonesia will experience higher growth rates. Nevertheless, their predicted average annual growth rates for the years 1992–2000 are predicted to decline gradually. China has experienced rapid economic growth since 1990, and is predicted to have an annual growth rate of roughly 12% up to the year 2000.

Finally, the improved grey model presented in this paper is not suitable for forecasting future traffic over very long periods of time due to the convergence problems of Markov-chain sign estimation. Forecasters may, however, regularly use new data to update the model to improve forecasting accuracy.

Acknowledgement

This work was supported by the National Science Council, R.O.C., through grant NSC 87-2211-E-009-012.

References

Behbehani, R. and Kanafani, A. (1980). *Demand and Supply Models of Air Traffic in International Markets.* Berkeley, Institute of Transportation Studies, Research Report ITS-WP-80-5.

Boeing (1993; 1995). *Trans-Pacific Passenger Market, U.S.–Europe Passenger Market, Europe-Orient Passenger Market.* Boeing Commercial Airplane Group, Seattle.

Boeing (1995). *World Market Outlook.* Boeing Commercial Airplane Group, Seattle.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control.* Prentice Hall, New Jersey.

Chen, C.K. and Tien, T.L. (1993). "Improved grey system identification in engineering system reliability-prediction." Keynote Paper of the 10th National Conference of the Chinese Society of Mechanical Engineers, Hsinchu, Taiwan.

Deng, J. (1982). *Grey System Fundamental Method.* Huazhong University of Science and Technology, Wuhan, China (in Chinese).

Deng, J. (1986). *Grey Prediction and Decision.* Huazhong University of Science and Technology, Wuhan, China (in Chinese).

Deng, J., Guo, H., Xu, S., Meng, W., Chen, M. and Xu, X. (1988). *Essential Topics on Grey System: Theory and Application.* A project supported by National Natural Science

Foundation of China, Huazhong University of Science and Technology, China Ocean Press, Beijing, China.

Deng, J. (ed) (1988). *Grey System Book*, Sci-Tech Information Services, Windsor.

Deng, J. (1989). "Introduction to grey system theory", *Journal of Grey System*, **1**(1), 1–24.

Federal Aviation Administration (1980). *FAA Aviation Forecasts, Fiscal Years 1981–1992*. Washington D.C.

Horonjeff, R. and McKelvey, X. (1994). *Planning and Design of Airports*. McGraw-Hill Inc., New York, N.Y.

Huth, W.L. and Eriksen, S.E. (1987). "Airline traffic forecasting using deterministic and stochastic time series decomposition." *Logistics and Transportation Review*, **23**(4), 401–409.

Ippolito, R.A. (1981). "Estimating airline demand with quality of service variables." *Journal of Transport Economics and Policy*, January, 7–15.

Kaemmerle, K.C. (1991). "Estimating the demand for small community air service." *Transportation Research*, **25A**(2/3), 101–112.

Nam, K. and Schaefer, T. (1995). "Forecasting international airline passenger traffic using neural networks." *Logistics and Transportation Review*, **31**(3), 239–252.

Russon, M.G. and Riley, N.F. (1993). "Airport substitution in short haul model of air transportation." *International Journal of Transportation Economics*, **XX**(2), 157–173.

Sun, G. (1991). "Prediction of vegetable yields by grey model GM(1,1)." *Journal of Grey System*, **2**, 187–197.

Young, K.H. (1972). " A synthesis of time-series and cross-section analysis: demand for air transportation service." *Journal of America Statistics Association*, **67**(339), 560–566.