Discussion Paper No. 533

TIME SERIES FORECASTS OF INTERNATIONAL TOURISM DEMAND FOR AUSTRALIA

Christine Lim
and
Michael McAleer

April 2001

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Time Series Forecasts of International Tourism Demand for Australia

Christine Lim

School of Tourism and Hotel Management

Griffith University

PMB 50 Gold Coast Mail Centre

Queensland 9726

Australia

Tel:	+617 5594 8824
Fax:	+617 5594 8507
E-mail:	Christine.Lim@mailbox.gu.edu.au

and

Michael McAleer

Department of Economics

University of Western Australia

Western Australia 6009

Australia

Time Series Forecasts of International Tourism Demand for Australia

Abstract

This paper examines stationary and nonstationary time series by formally testing for the presence of unit roots and seasonal unit roots prior to estimation, model selection and forecasting. Various Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) models are estimated over the period 1975(1)-1989(4) for tourist arrivals to Australia from Hong Kong, Malaysia and Singapore. The mean absolute percentage error (MAPE) and root mean squared error (RMSE) are used as measures of forecast accuracy. As the best fitting ARIMA model is found to have the lowest RMSE, it is used to obtain post-sample forecasts. Tourist arrivals data for 1990(1) to 1996(4) are compared with the forecast performance of the ARIMA model for each origin market. The fitted ARIMA model forecasts tourist arrivals from Singapore between 1990(1)-1996(4) very well. Although the ARIMA model outperforms the seasonal ARIMA models for Hong Kong and Malaysia, the forecast of tourist arrivals is not as accurate as in the case of Singapore.

Keywords: Unit roots, seasonality, forecasting models, forecast accuracy, root mean square error, ex post forecasts.

1. Introduction

Hong Kong, Malaysia and Singapore are Australia's major tourist markets in Asia apart from Japan. In terms of the international tourism market share of the three countries, Singapore is Australia's fifth major market, with Hong Kong and Malaysia occupying seventh and eighth places, respectively. The average annual growth rates of tourist arrivals from Hong Kong, Malaysia and Singapore during 1991-96 were 19.37%, 21.4% and 20.8%, respectively, which increased from average growth rates of 17.7%, 9.7% and 16.1% over the period 1985-90. The rise in inbound tourism from these markets could be attributed to the rapid economic growth experienced by these countries in the first half of the 1990s. Inspite of the phenomenal growth of inbound tourism from these source markets, which far exceeded the average growth rate of international arrivals to Australia of 10.5% over the period 1990-96, little research on these markets has been undertaken so as to understand their significant contributions to Australia's inbound tourism. Most of the research has been conducted on the four major markets to Australia, namely Japan, New Zealand, United Kingdom and the United States.

Quantitative methods for generating forecasts of future outcomes using statistical procedures involve the examination of current and historical seasonally unadjusted data. This knowledge can be used to extrapolate the variable of interest. It is assumed that the process is stable over the forecast time horizon, but this assumption may only be valid for short-term forecasts. Two types of quantitative forecasting models used are time series models and causal econometric models. Time series models involve a statistical analysis which uses only the historical data of the variable to be forecast. Causal models are based on the statistical analysis of data for other related (explanatory) variables, and the use of these variables to forecast the dependent variable of interest.

At present there are numerous forecasting methods available and the empirical findings, which are often in conflict, have given no clear guidelines as to the most appropriate methods for forecasting. The literature on international tourism demand forecasting, based on different univariate time series forecasting methods (see, for example, Geurts and Ibrahim, 1975; Choy, 1984; van Doorn, 1984; Martin and Witt, 1989; Chan, 1993; Witt et al., 1994; Turner et al., 1995, 1997; Frechtling, 1996; Kulendran and King, 1997; Chu, 1998; Kim, 1999), is numerous.

Forecasting performances of the various models are affected by the type of data used (namely monthly, quarterly or annual data), the forecasting horizon, and the country of origin. From some these studies, it is clear that sophisticated procedures such as ARIMA models do not necessarily forecast better than their simple counterparts.

There are a number of factors used to evaluate the effectiveness of a forecasting method, such as forecasting accuracy, costs associated with the application of a forecasting procedure (for example, installation and operating costs), and ease of application and interpretation of the output from a forecasting method. Accuracy is often regarded as the dominant criterion for selecting a forecasting method. The accuracy of a forecasting method is determined by analyzing the forecast error, which is defined as the actual value minus the forecast (or fitted) value of the variable for time period t, namely:

$$e_t = A_t - F_t$$

where

e_t = forecast error at time t;

A_t = actual tourist arrivals at time t;

F_t = forecast tourist arrivals at time t+1.

For instance, forecast optimization typically chooses a model that minimizes root mean squared error (RMSE), which is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}\ .$$

In this paper, various Box-Jenkins (1970) autoregressive integrated moving average (ARIMA) forecasting models are considered and their comparative performances analysed over a sample of international tourism demand for Australia by each of three origin countries, namely Hong Kong, Malaysia and Singapore. The ARIMA models provide a useful framework to understand how the tourism time series is generated. Unlike univariate smoothing models which are more commonly used, the ARIMA approach requires a tourism time series to be tested for nonstationarity prior to undertaking estimation and forecasting exercise. If a series is nonstationary (that is, the series has a mean and variance that are not constant over time), the series has to be differenced to transform it to a stationary series, before generating forecasts. A stationary tourist arrival series typically provides better and more reliable forecasts. Very few of the recent published studies on tourism forecasting have considered or presented tests for unit roots and seasonal unit roots before estimating ARIMA models and using them for forecasting. Such tests and their implications will be discussed in this paper for the historical data on individual tourists arrivals to Australia from Hong Kong, Malaysia and Singapore.

The logarithms of quarterly tourist arrivals data (from the Australian Bureau of Statistics) for the March quarter of 1975 to the December quarter of 1989 are used. One-quarter-ahead international tourist arrivals forecasting accuracy, beyond the sample used for estimation, is evaluated for the period 1990(1)-1996(4) using various estimated ARIMA time series models. The motivation for the choice of this period is twofold. First, between 1990 and 1996, Australia experienced the largest average annual percentage growth in tourist arrivals from Asia of 23 percent. Comparison of the out-of-sample forecast of the various ARIMA models would be considered useful for a wide range of policy-making in the tourism and travel industry. Second, the significance of the impact on international tourist arrivals to Australia due to the 1979 Oil Crisis, 1988 Bicentennial Celebration and 1989 Air Pilots Dispute can be estimated.

2. ARIMA Forecasting Models

The logarithms of quarterly tourist arrivals from the three countries are used to capture the multiplicative effects in the levels of the variables. Using an autoregressive specification and ordinary least squares estimation, current tourist arrivals can be forecast one quarter ahead, based on a fourth-order process, as follows:

$$A_t = \beta_0 + \beta_1 A_{t-1} + \beta_2 A_{t-2} + \beta_3 A_{t-3} + \beta_4 A_{t-4} + \varepsilon_t .$$

Table 1 shows that the only significant lags in forecasting tourist arrivals are the second and fourth for Hong Kong, the first and fourth lags for Malaysia, and only the fourth lag for Singapore.

The influential work of Box and Jenkins (1970) shifted professional attention in time series modelling away from stationary processes to a class of nonstationary processes and the related ideas of the order of integration necessary to obtain stationary series. Furthermore, the Box-Jenkins method is popular because of its generality since it can handle any stationary or nonstationary time series, with and without seasonal elements. The frequent use in applied empirical work and availability in well-documented econometric computer programs have perhaps contributed most to its popularity. The Box-Jenkins method for selecting an appropriate autoregressive integrated moving average (ARIMA) model for estimating and forecasting a univariate time series consists of identification, estimation and testing, and application. In the identification phase, a general class of models applicable to a particular situation is examined with the aid of the sample correlograms, and autocorrelation and partial autocorrelation functions.

The original time series in logarithms are checked for stationarity using the augmented Dickey-Fuller (ADF) test for unit roots and, if necessary, the series are transformed by taking appropriate differences to render the series stationary. A detailed explanation of the test procedure is given in Lim and McAleer (2000a). The ADF tests which are performed sequentially show that the fourth lag is significant, and the ADF test statistics with trend are –2.40, -1.83 and –2.52 for Hong Kong, Malaysia and Singapore, respectively. Each of the calculated statistics exceeds the critical value of –3.49 at the 5% significance value, so the null hypothesis of a unit root is not rejected, which implies that each of the three tourist arrival series is nonstationary. Taking first differences renders each series stationary, with the ADF statistics in all cases for lag length of three (that is, -3.77, -4.40 and –4.74 for Hong Kong, Malaysia and Singapore, respectively) being less than the critical value of –2.91 at the 5% significance level.

Since tourism data exhibit varying seasonal patterns, it is imperative to test for the presence of seasonal unit roots in univariate series. The test most often used is the HEGY test of Hylleberg, Engle, Granger and Yoo (1990). A detailed explanation of the test procedure is given, for example, in Lim and McAleer (2000b). Briefly, the relevant hypotheses to be tested are as follows:

1) $\pi_1 = 0$; unit root at the zero frequency

2) $\pi_2 = 0$; unit root at the semi-annual frequency

3) $\pi_3 = \pi_4 = 0$; unit root at the annual frequency.

The results (see Table 2) indicate that quarterly tourist arrivals from Hong Kong, Malaysia and Singapore are each integrated at the zero and semi-annual frequencies, but not at the annual frequency, I(1,1,0).

Subsequently, parameter estimation in phase two involves fitting various autoregressive (AR) models of order p and moving average (MA) models of order q. The best fitting parsimonious model is selected, based on various criteria, such as statistically significant AR and MA estimated coefficients at the 5% level, absence of serial correlation, and optimisation of the Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC). Diagnostic checking involves residual analysis to ensure that the estimated model has independent and identically distributed errors. When a satisfactory model has been selected, it is used in the third phase to forecast future values of tourist arrivals from the three origin countries. Using least squares estimation, the variance of the optimum i-period-ahead forecast will be less than the variance of the predicted series:

$$\mathrm{Var}(A_{t+i}) = \sigma_A^2 = \mathrm{Var}(F_{t+i}) + \sigma_\varepsilon^2 .$$

When an ARIMA model has been fitted to a time series, the i-period-ahead forecast of tourist arrivals is given by:

$$F_{t+i} = \hat{C} + \hat{\phi}_1 \hat{A}_{t+i-1} + \ldots + \hat{\phi}_{p+d} \hat{A}_{t+i-p-d} + \hat{\varepsilon}_{t+i} - \hat{\theta}_1 \hat{\varepsilon}_{t+i-1} - \ldots - \hat{\theta}_q \hat{\varepsilon}_{t+i-q} , \quad i = 1, 2, \ldots \qquad (8)$$

In order to compare the ex post forecast accuracy of the various ARIMA models for 1990(1)-1996(4), the best fitting ARIMA models are estimated separately for tourist arrivals series from 1975(1) to 1989(4). The correlogram and unit root tests of the series before and, if necessary, after differencing are examined for stationarity. After empirical examination, the most appropriate models for tourist arrivals from Hong Kong, Malaysia and Singapore are determined as ARIMA(3,1,1), ARIMA(3,1,2) and ARIMA(4,1,0), respectively (with absolute t-ratios in parentheses):

$$(1 + 1.32L + 0.82L^2 + 0.52L^3)(1 - L)\log HK_t = 0.04 + \hat{\varepsilon}_{HKt} - 0.97\hat{\varepsilon}_{HK(t-1)}$$

(10.7) (4.20) (4.15) (3.25) (38.4)

AIC = -3.65, SBC = -3.47 (Hong Kong)

$$(1 + 0.69L + 0.98L^2 + 0.65L^3)(1 - L)\log M_t = 0.03 + \hat{\varepsilon}_{Mt} - 0.95\hat{\varepsilon}_{M(t-1)}$$

(6.68) (117) (6.52) (2.41) (27.9)

AIC = -3.73, SBC = -3.55 (Malaysia)

$$(1+0.64L+0.58L^2+0.6L^3-0.36L^4)(1-L)\log S_t = 0.04+\hat{\varepsilon}_{St}$$
$$\quad (4.84) \quad (4.26) \quad (4.45) \quad (2.67) \qquad\qquad\qquad (4.94)$$

AIC = -3.95, SBC = -3.77. (Singapore)

Since the specific ARIMA models that adequately describe tourist arrivals from Hong Kong, Malaysia and Singapore are given above, the fitted models used for ex post forecasting tourist arrivals from these origin countries are given as follows:

$$\begin{aligned} F_{\log HK(t+i)} &= 0.04 - 0.32\log HK_{t+i-1} + 0.50\log HK_{t+i-2} + 0.30\log HK_{t+i-3} \\ &\quad + 0.52\log HK_{t+i-4} + \hat{\varepsilon}_{HK(t+i)} - 0.97\hat{\varepsilon}_{HK(t+i-1)} \end{aligned} \qquad \text{(Hong Kong)}$$

$$\begin{aligned} F_{\log M(t+i)} &= 0.03 + 0.31\log M_{t+i-1} - 0.29\log M_{t+i-2} + 0.33\log M_{t+i-3} \\ &\quad + 0.65\log M_{t+i-4} + \hat{\varepsilon}_{M(t+i)} - 0.95\hat{\varepsilon}_{M(t+i-1)} \end{aligned} \qquad \text{(Malaysia)}$$

$$\begin{aligned} F_{\log S(t+i)} &= 0.04 + 0.36\log S_{t+i-1} + 0.06\log S_{t+i-2} - 0.02\log S_{t+i-3} \\ &\quad + 0.96\log S_{t+i-4} - 0.36\log S_{t+i-5} + \hat{\varepsilon}_{S(t+i)} \end{aligned} \qquad \text{(Singapore)}$$

where t = 1989(4), i = 1, 2, ...

The expected values of the future random errors are assumed to be zero.

As the individual quarterly international tourist arrivals to Australia from Hong Kong, Malaysia and Singapore exhibit pronounced seasonality, the most appropriate multiplicative seasonal ARIMA models selected for the various tourist arrivals series can be used to forecast future observations, namely ARIMA(1,1,4)(0,1,0)$_4$ for Hong Kong, ARIMA(2,1,1)(4,1,2)$_4$ for Malaysia, and ARIMA(0,1,4)(0,1,1)$_4$ for Singapore, which are given as follows:

$$(1+0.48L)(1-L)(1-L^4)\log HK_t = (1+0.93L^4)\hat{\varepsilon}_{HKt} \qquad \text{(Hong Kong)}$$
$$\quad (4.21) \qquad\qquad\qquad\qquad (39.2)$$

AIC = -3.98, SBC = -3.91

$$(1-0.62L^2)(1+0.42L^4)(1-L)(1-L^4)\log M_t = (1+0.27L)(1+0.98L^2)\hat{\varepsilon}_{Mt} \qquad \text{(Malaysia)}$$

$$\quad (6.81) \qquad (3.06) \qquad\qquad\qquad\qquad (2.81) \qquad (4746)$$

AIC = -3.76, SBC = -3.61

$$(1-L)(1-L^4)\log S_t = (1+0.62L^4)(1+0.56L)\hat{\varepsilon}_{St} \qquad \text{(Singapore)}$$

$$\qquad\qquad\qquad (5.88) \qquad (5.06)$$

AIC = -4.17, SBC = -4.10.

3. Magnitude of Forecasting Errors

With the final observation being tourist arrivals for the fourth quarter of 1989, Table 3 presents the RMSE one-quarter-ahead forecast accuracy measure of the ARIMA and Multiplicative Seasonal ARIMA models. For tourist arrivals from Hong Kong and Malaysia, the ARIMA model forecasts better than the seasonal multiplicative ARIMA model, and the reverse holds for Singapore. However, the mean absolute percentage error (MAPE) of the ARIMA model is lower than that of the seasonal model for tourist arrivals from Singapore.

Using the best fitting model for each tourist arrivals series, the accuracy of post-sample forecasts from 1990(1) to 1996(4) is obtained to examine the relationship between the fit of the model and forecast performance. Table 4 shows Theil's (1966) U-coefficients for the ARIMA and seasonal ARIMA forecasts, all of which are less than one. As the U-coefficients of the ARIMA model are all less than that of the seasonal ARIMA for each series, these results suggest that the ARIMA model performs better in forecasting tourist arrivals from the three origin countries for the period 1990(1) to 1996(4). Using the best fitting ARIMA model for post-sample forecasting, the sample coefficient of correlation is computed as a goodness-of-fit measure to determine how well the actual values fit the future observations. This measure provides useful information as to how well the model forecasts the data. Table 5 shows that the correlation coefficients of the ARIMA model range from 0.67 to 0.95. Apparently, the fitted ARIMA model forecasts tourist arrivals from Singapore very well, given that 95% of the variation in the tourist arrivals forecast is associated with variations in actual tourist arrivals between 1990(1)-1996(4). Even though the ARIMA model outperforms the other models in forecasting tourist arrivals from Malaysia, only 67% of the variation in the tourist arrivals forecast is associated with variations in the actual tourist arrivals in the same period.

The fitted values, which are interpreted as the forecasts for the next quarter, are sufficiently close to the actual values for tourist arrivals from Hong Kong, Malaysia and Singapore using the ARIMA models, as shown in Figures 1-3. However, the forecasts from the seasonal ARIMA models are not close to the actual values (these figures are available on request). There is no systematic pattern in the residuals for the ARIMA models. Using 24 lags for each series, the correlograms of the residuals show that there are very few autocorrelations outside the bounds ±0.3, or the 95% confidence interval. The autocorrelations which are not within two standard errors of the mean include the sixth and tenth lags for Hong Kong, the eleventh lag for Malaysia, and only marginally for the third, sixth and tenth lags for Singapore (these figures are also available on request). The Lagrange multiplier test for serial correlation, LM(SC), shows that the null hypothesis of no serial correlation is not rejected, and hence serial correlation is deemed to be absent in the residuals (see Table 6). The calculated F values, which lie between 0.83 and 3.03, are all less than the critical F value at the 5% level.

Table 7 shows the sum of squared residuals (SSR) from the fitted ARIMA and seasonal ARIMA models for Hong Kong, Malaysia and Singapore: $SSR = \sum_{t=1}^{n} \hat{e}_t^2$, where $\hat{e}_t$ denotes the one-quarter-ahead prediction error at time t. The standard deviation (SD) of residuals, calculated as $\sqrt{\hat{\sigma}_e^2} = \sqrt{SSE / n}$, is between 0.12 and 0.15, signifying the average deviations in one-quarter-ahead forecasts of tourist arrivals.

Table 8 shows a breakdown of RMSE for the fitted ARIMA and seasonal ARIMA models according to the forecasting horizon for the three tourist arrivals series. It suggests that a shorter forecasting horizon is more accurate than a longer forecasting horizon for tourist arrivals from Malaysia and Singapore. However, the same argument does not hold according to the RMSE for tourist arrivals from Hong Kong.

During the period 1975(1) to 1989(4), Australia experienced the Oil Price Crisis of 1979, the Bicentennial Celebration in 1988 of European Settlement in Australia, and the Australian Air Pilots Strike in1989-90. These one-off events could distort the estimation, testing and analysis of the underlying process, which is critical in forecasting. In order to analyse the impact of these one-off events, intervention analysis with deterministic dummies allows the effects of these exogenous shocks to be represented by dynamic ARIMA models, as follows (with absolute t-ratios in parentheses):

$$\underset{}{(1+\underset{(10.9)}{1.36}L+\underset{(5.23)}{0.87}L^2+\underset{(6.31)}{0.55}L^3)(1-L)\log HK_t} = \underset{(3.84)}{0.04}+\underset{(1.48)}{0.07}D1-\underset{(0.92)}{0.04}D2+\underset{(1.19)}{0.11}D3+\hat{\varepsilon}_{HKt}-\underset{(12.6)}{1.13}\hat{\varepsilon}_{HK(t-1)} \quad \text{(Hong Kong)}$$

$$(1+\underset{(6.55)}{0.7}L+\underset{(112)}{0.98}L^2+\underset{(6.40)}{0.66}L^3)(1-L)\log M_t = \underset{(2.28)}{0.03}+\underset{(0.49)}{0.02}D1-\underset{(0.73)}{0.04}D2-\underset{(0.51)}{0.03}D3+\hat{\varepsilon}_{Mt}-\underset{(26.6)}{0.96}\hat{\varepsilon}_{M(t-1)} \quad \text{(Malaysia)}$$

$$(1+\underset{(4.89)}{0.69}L+\underset{(4.36)}{0.64}L^2+\underset{(4.56)}{0.66}L^3-\underset{(2.10)}{0.3}L^4)(1-L)\log S_t = \underset{(4.70)}{0.04}+\underset{(1.08)}{0.03}D1-\underset{(0.96)}{0.03}D2-\underset{(0.09)}{0.01}D3+\hat{\varepsilon}_{St} \quad \text{(Singapore)}$$

where

D1 = dummy variable for the 1979 Oil Price Crisis;

D2 = dummy variable for the 1988 Bicentennial Celebration;

D3 = dummy variable for the 1989-90 Air Pilots Strike.

The impulse specification characterizes a temporary intervention, in which D1, D2 and D3 are zero for all periods except for the quarters in which the events occurred. These include impulse (or dummy) variables for the Oil Price Crisis for the period 1979(1)-1979(4), Bicentennial Celebration for 1988(1)-1988(4) and Air Pilots Dispute for 1989(3)-1990(2). Witt et al. (1994) obtained mixed results when their ARIMA models incorporated interventions. However, Table 9 shows that none of the intervention variables used in this study is significant at the 5% level. In addition, the constant term, and the autoregressive and moving average coefficients remain significant at the 5% level.

4. Conclusion

This paper examines univariate time series ARIMA forecasting methods based on current and past tourist arrivals from three Asian countries to Australia. A wide range of quantitative forecasting techniques is available, from sophisticated regression and smoothing procedures to naive models. Unlike many exponential smoothing procedures which attempt to fit the data to a particular model, time series analysis of ARIMA models fits various models to historical data to obtain forecasts of tourist arrivals from Hong Kong, Malaysia and Singapore. Makridakis and Hibon (1979) reported little or no improvement in empirical forecast accuracy from using ARIMA models instead of simpler forecasting techniques, which seems to support Geurts and Ibrahim's (1975, p. 186) contention that: "... the model that fits best is not necessarily the one that forecasts best".

Unlike the numerous studies using the Box-Jenkins model, the following is undertaken in this paper prior to computing forecast accuracy measures for both levels and logarithms:

- pre-testing for nonstationarity in the series by applying the augmented Dickey-Fuller tests for unit roots, and
- model selection of the best fitting ARIMA model (using various criteria) to explain the pattern of tourist arrivals based on its preceding values, and current and previous random errors for each origin or tourist-source country.

The use of such procedures, particularly tests for unit roots, improves the validity of using the ARIMA models for forecasting and allows the forecaster to make informed judgments at each step as the results are presented by the statistical packages. Overall, this paper shows that by comparing the root mean squared errors, lower post-sample forecast errors were obtained when time series methods, such as the Box-Jenkins ARIMA and seasonal ARIMA models, were used.

The Box-Jenkins approach is often argued to be powerful but also complex to use in building forecasting models. Besides forecasting accuracy, simplicity of technique and the cost aspects of various techniques (which include labour skills, financial means, and time) are often stated as being useful criteria considered by policy-makers and forecasters. Admittedly, in the Box-Jenkins approach, the initial model selection stage encounters problems of practical implementation, and it can be time consuming if a large number of time series observations are to be analysed. Even if van Doorn (1984, p. 25) is correct in the observation that: "the popularity and perceived usefulness of a technique is directly related to the effort and sophistication required to implement that technique", the appropriateness of the use of the Box-Jenkins approach for tourism practitioners should not be underestimated in the art of forecasting international tourist flows to a particular destination. The Box-Jenkins approach is flexible and a broad class of general models can be considered. Diagnostic tests are also used in this study to test the validity of the ARIMA models selected, for example, the Lagrange multiplier test for serial correlation. The lack of such diagnostic checks in previous studies suggests that inferences from estimated models may be highly sensitive to the assumption of serial independence. This paper has illustrated the challenges that tourism planners, policy-makers and academics will encounter if ARIMA models are to be used sensibly in building forecasting models.

Acknowledgements

The first author wishes to acknowledge the financial support of an Individual Research Grant, Faculties of Economics & Commerce, Education and Law, University of Western Australia, and the Australian Research Council. The second author wishes to acknowledge the financial support of the Australian Research Council. This paper was revised while the second author was visiting the Institute of Social and Economic Research at Osaka University.

References

Australian Bureau of Statistics. Various Issues of Short-Term Overseas Arrivals, 1975-1996, Canberra.

Box, G.E.P. and G.M. Jenkins (1970). *Time Series Analysis: Forecasting and Control*, Holden Day, San Francisco.

Chan, Y.-M. (1993). Forecasting Tourism: A Sine Wave Time Series Regression Approach. *Journal of Travel Research*, 32, 58-60.

Choy, D. (1984). Forecasting Tourism Revisited. *Tourism Management*, 5, 171-176.

Chu, F.-L. (1998). Forecasting Tourist Arrivals. *Journal of Travel Research*, 36, 79-84.

Frechtling, D.C. (1996). *Practical Tourism Forecasting*, Butterworth-Heinemann, Oxford.

Geurts, M. and I. Ibrahim (1975). Comparing the Box-Jenkins Approach with the Exponentially Smoothed Forecasting Model Application to Hawaii Tourists. *Journal of Marketing Research*, 12, 182-188.

Hylleberg, S., R.F. Engle, C.W.J. Granger and B.S. Yoo (1990). Seasonal Integration and Cointegration. *Journal of Econometrics*, 99, 215-238.

Kim, J.H. (1999). Forecasting Monthly Tourist Departures from Australia. *Tourism Economics*, 5, 277-291.

Kulendran, N. and M.L. King (1997). Forecasting International Quarterly Tourist Flows using Error-Correction and Time-Series Models. *International Journal of Forecasting*, 13, 319-327.

Lim, C. and M. McAleer (2000a). Monthly Seasonal Variations: Asian Tourism to Australia. *Annals of Tourism Research*, 28, 68-82.

Lim, C. and M. McAleer (2000b). A Seasonal Analysis of Asian Tourist Arrivals to Australia. *Applied Economics*, 32, 499-509.

Makridakis, S. and M. Hibon (1979). Accuracy of Forecasting: An Empirical Investigation. *Journal of the Royal Statistical Society A*, 142, 97-145.

Martin, C.A. and S.F. Witt (1989). Forecasting Tourism Demand: A Comparison of the Accuracy of Several Quantitative Methods. *International Journal of Forecasting*, 5, 7-19.

Turner L., N. Kulendran and V. Pergat (1995). Forecasting New Zealand Tourism Demand with Disaggregated Data. *Tourism Economics*, 1, 51-69.

Turner, L.W., N. Kulendran and H. Fernando (1997). Univariate Modelling Using Periodic and Non-Periodic Analysis: Inbound Tourism to Japan, Australia and New Zealand compared. *Tourism Economics*, 3, 39-56.

Theil, H. (1966). *Applied Economic Forecasting*, North Holland, Amsterdam.

van Doorn, J.W.M. (1984). Tourism Forecasting and the Policymaker: Criteria of Usefulness. *Tourism Management*, 5, 24-39.

Witt, C.A., S.F. Witt and N. Wilson (1994). Forecasting International Tourist Flows. *Annals of Tourism Research*, 21, 612-628.

Table 1

Autoregressions of the Logarithm of
International Tourist Arrivals, 1975(1)-1989(4)

Variable	Hong Kong	Malaysia	Singapore
Constant	-0.01 (-0.04)	0.36 (1.04)	0.08 (0.36)
A_{t-1}	0.20 (1.72)	0.19 (2.16)	0.04 (0.71)
A_{t-2}	0.32 (2.76)	-0.07 (-0.78)	0.05 (0.99)
A_{t-3}	-0.11(-0.91)	0.13 (1.46)	-0.02 (-0.40)
A_{t-4}	0.60 (5.09)	0.72 (8.86)	0.94 (17.6)

Note: t-statistics are given in parentheses.

Table 2

HEGY Tests for Seasonal Integration of Quarterly Tourist Arrivals

Origin Country	$t(\pi_1)$	$t(\pi_2)$	$F(\pi_3, \pi_4)$
Hong Kong	-2.65	4.38	23.05
Malaysia	-1.46	4.06	10.73
Singapore	-2.35	2.24	15.45

Note: An intercept, three seasonal dummies and a time trend are included in the HEGY regressions. n = 84 is the number of observations in each series. The critical values at the 5% level are taken from Hylleberg et al. (1990) for 100 observations: $t(\pi_1)$ = -3.53, $t(\pi_2)$ = -2.94, and $F(\pi_3, \pi_4)$ = 6.60.

Table 3

RMSE for One-Quarter-Ahead Ex Post Forecasts of the Logarithm of International Tourist Arrivals, 1975(1)-1989(4)

Origin Country	ARIMA	Seasonal ARIMA
Hong Kong	0.016 (1)	0.042 (2)
Malaysia	0.028 (1)	0.061 (2)
Singapore	0.035 (2)	0.008 (1)

Note: Figures in parentheses denote rankings.

Table 4

Theil's U-Coefficients for Forecast Errors, 1990(1)-1996(4)

Origin Country	ARIMA	Seasonal ARIMA
Hong Kong	0.468	0.717
Malaysia	0.439	0.728
Singapore	0.090	0.885

Table 5

Correlation Coefficients Between Actual and Predicted Values
Using Box-Jenkins ARIMA Models, 1990(1)-1996(4)

Origin Country	ARIMA	Correlation Coefficient
Hong Kong	(3,1,1)	0.68
Malaysia	(3,1,2)	0.67
Singapore	(4,1,0)	0.95

Table 6

Lagrange Multiplier Test for Serial Correlation

Origin Country	Model	F value	p-value
Hong Kong	ARIMA(3,1,1)	3.03	0.06
Malaysia	ARIMA(3,1,2)	1.38	0.26
Singapore	ARIMA(4,1,0)	0.83	0.44

Note: The LM test statistics presented here are for second-order serial correlation. Test results for fourth-order serial correlation were qualitatively very similar.

Table 7

Sum of Squared Residuals (SSR) and Standard Deviation (SD) of Residuals for ARIMA and Seasonal ARIMA Models

Origin Country	SSR		SD	
	ARIMA	Seasonal ARIMA	ARIMA	Seasonal ARIMA
Hong Kong	1.21	0.93	0.15 (56)	0.13 (54)
Malaysia	1.13	0.96	0.14 (56)	0.14 (49)
Singapore	0.88	0.79	0.13 (55)	0.12 (55)

Note: Sample sizes are given in parentheses.

Table 8

Forecast RMSE for Different Horizons

Forecast Horizon (quarters ahead)	Hong Kong	Malaysia	Singapore
1	0.016	0.03	0.03
2	0.03	0.07	0.03
3	0.59	0.20	0.01
4	0.44	0.08	0.02
5	0.13	0.009	0.05
6	0.02	0.10	0.03
7	0.05	0.17	0.007
8	0.07	0.002	0.01
RMSE: up to 4 quarters ahead	0.439	0.126	0.033
RMSE: 5 to 8 quarters ahead	0.073	0.151	0.038
RMSE: 1990(1)-1996(4)	0.362	0.169	0.096

Table 9

Impact Effects on International Tourist Arrivals to Australia of Interventions in 1979, 1988 and 1989

Origin Country	Oil Crisis 1979	Bicentennial Celebration 1988	Air Pilots Dispute 1989
Hong Kong	0.07	-0.04	0.11
	(1.48)	(-0.92)	(1.19)
Malaysia	0.02	-0.04	-0.03
	(0.49)	(-0.73)	(-0.51)
Singapore	0.03	-0.03	-0.006
	(1.08)	(-0.96)	(-0.09)

Note: t-statistics are given in parentheses.

Figure 1

Actual, Fitted and Residuals from ARIMA(3,1,1) Model
of Tourist Arrivals from Hong Kong, 1975(1)-1989(4)

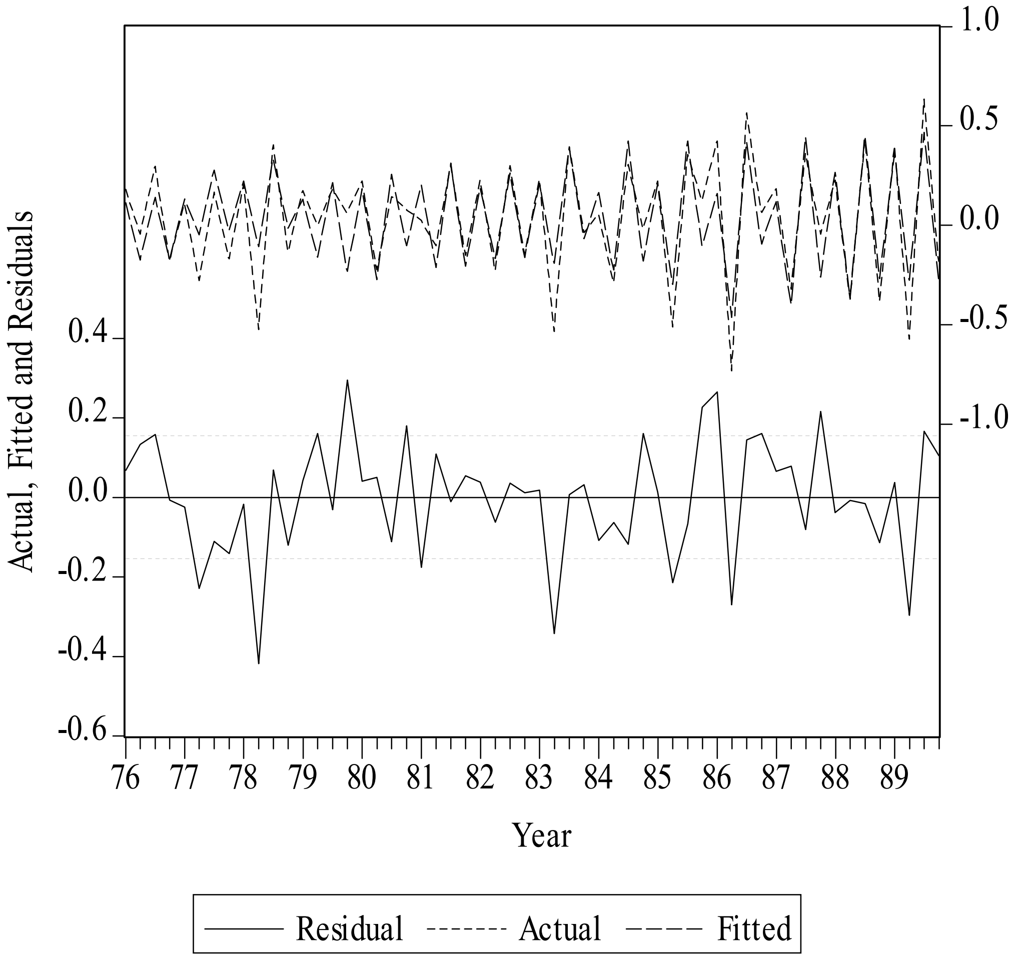

Figure 2

Actual, Fitted and Residuals from ARIMA(3,1,2) Model of Tourist Arrivals from Malaysia, 1975(1)-1989(4)

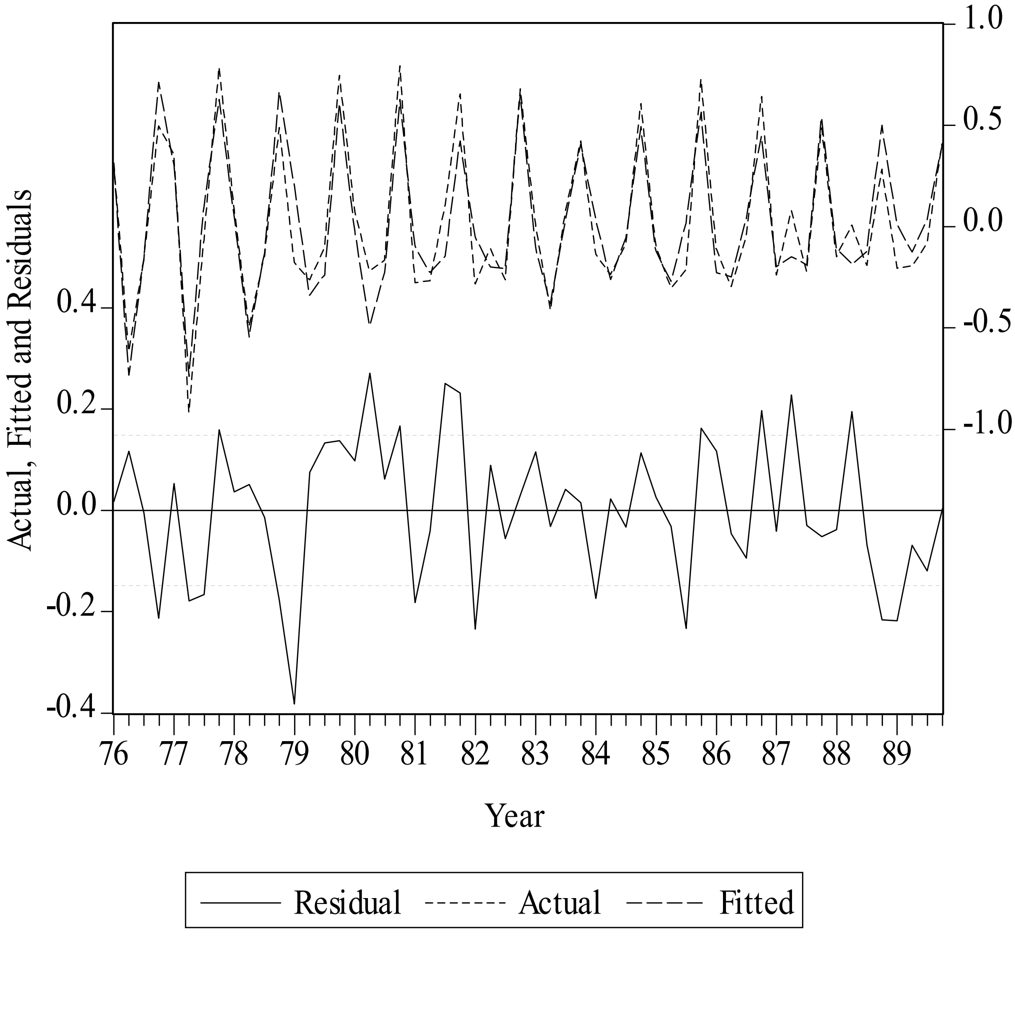

Figure 3

Actual, Fitted and Residuals from ARIMA(4,1,0) Model of Tourist Arrivals from Singapore, 1975(1)-1989(4)