Forecasting Quarterly Tourism Demand by Hong Kong for Australia

Christine Lim

School of Finance and Business Economics
Edith Cowan University

Michael McAleer

Department of Economics
University of Western Australia

Abstract: The purpose of this paper is to forecast tourism demand by Hong Kong for Australia. Various forecasting models are estimated over the period 1975(1)-1989(4) for seasonally unadjusted quarterly tourist arrivals to Australia from Hong Kong. As the best fitting Autoregressive Integrated Moving Average (ARIMA) and seasonal ARIMA models outperform a variety of simple Smoothing models, the time series models are used to obtain the post-sample forecasts, with root mean squared error being used to evaluate their accuracy.

1. Introduction

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

t = 1, 2, ... , n.

For a given data process and method, the forecast error is assumed to be an independent random variable with zero mean and constant variance: $E(e_t) = 0$ and $Var(e_t) = \sigma_e^2$. Although accuracy is inversely related to the size of the forecast error, there is not a universally accepted measure of such accuracy. A variety of measures of accuracy is available, including the root mean squared error (RMSE) and the mean absolute percentage error (MAPE) of the forecasts. Forecast optimization typically chooses a model that minimizes RMSE, which is calculated as:

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

In this paper, various univariate forecasting methods are considered and their comparative performance is analysed with regard to international tourism demand by Hong Kong for Australia. The econometric software program EViews provides several exponential smoothing methods which give one-period-ahead forecasts. 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. Using single-equation models, one-quarter-ahead international tourist arrivals forecast accuracy is evaluated for various procedures, including the smoothing and Box-Jenkins models, namely single exponential, Brown's double exponential, Holt-Winters seasonal and non-seasonal exponential smoothing, and the best fitting ARIMA time series models.

2. Exponential Smoothing and ARIMA Forecasting Models

Using the single exponential smoothing method, the forecast for t+i is given by the exponentially smoothed tourist arrivals series in t, which is the weighted average of the actual tourist arrivals in t and the smoothed lag series (which is the forecast made in the previous period). A single exponential smoothing method, which produces an i-period-ahead forecast (F_{t+i}) at time t, can be calculated recursively as follows:

$$\begin{aligned} F_{t+i} &= \tilde{A}_t = \alpha A_t + (1-\alpha)\tilde{A}_{t-1} \\ &= \alpha A_t + (1-\alpha)F_{t+i-1} \end{aligned},$$
$$0<\alpha<1, \qquad i = 1, 2, \ldots \qquad (1)$$

where

A_t = actual tourist arrivals at time t;

$\tilde{A}_t$ = smoothed estimate of tourist arrivals at time t;

F_{t+i} = single exponential smoothing forecast of tourist arrivals at time t+i;

α = smoothing coefficient.

Forecasts are given by the latest available smoothed value such that, at time t, the previous smoothed estimate, $\tilde{A}_{t-1}$, is updated as the new observation, A_t, becomes available. Thus, the new smoothed estimate, $\tilde{A}_t$, is the weighted average of A_t and $\tilde{A}_{t-1}$. The forecast in period t+i is based on weighting the observation in period t by a smoothing coefficient, α, and the most recent forecast by (1-α), so that F_{t+i} is a weighted average of all current and past observations.

Equation (1) can be rewritten as:

$$F_{t+i} = F_{t+i-1} + \alpha(A_t - F_{t+i-1}),$$

or, for i = 1,

$$F_{t+1} = F_t + \alpha(A_t - F_t) = F_t + \alpha e_t,$$

from which it can be seen that the single exponential smoothing procedure adjusts the forecast by a proportion of the most recent forecast error: if α is close to one, the new forecast is the previous forecast plus a substantial proportion of the most recent forecast error; if α is close to zero, the new forecast is very close to the previous forecast, with little influence from the most recent forecast error.

Single exponential smoothing is appropriate only for stationary and non-seasonal time series data with no structural change. As the international tourist arrivals series from Hong Kong exhibit clear trends, Brown's (1963) double exponential smoothing, which includes a deterministic trend in the forecast, could also be useful. Initially, the observed international tourist arrivals time series are smoothed by adding a weighted estimate of the smoothed lag to the latest observation, A_t, thereby yielding the smoothed series $\tilde{A}_t^{(1)}$. The latter can be smoothed again using the same coefficient, yielding the doubly smoothed series $\tilde{A}_t^{(2)}$, that is:

$$\tilde{A}_t^{(1)} = \alpha A_t + (1-\alpha)\tilde{A}_{t-1}^{(1)} \qquad (2a)$$

and

$$\tilde{A}_t^{(2)} = \alpha \tilde{A}_t^{(1)} + (1-\alpha)\tilde{A}_{t-1}^{(2)} \qquad (2b)$$

where $\tilde{A}_t^{(1)}$ and $\tilde{A}_t^{(2)}$ are the single and double exponential smoothing estimates of tourist arrivals, respectively. Brown's double exponential smoothing i-period-ahead forecast of international tourist arrivals is given by a constant term plus a linear trend term (T):

$$F_{t+i} = \tilde{A}_t + iT_t, \qquad i = 1, 2, \ldots \qquad (3)$$

where

$$\tilde{A}_t = 2\tilde{A}_t^{(1)} - \tilde{A}_t^{(2)}$$

and

$$T_t = (\alpha/(1-\alpha))(\tilde{A}_t^{(1)} - \tilde{A}_t^{(2)}).$$

Unlike the single exponential smoothing procedure, which gives a forecast that ignores the trend for all future observations, the double smoothing method produces a forecast that changes with a trend. The smoothing coefficient, α, is determined to minimize the forecast RMSE.

Brown's double exponential smoothing method only requires one smoothing coefficient to be chosen. The Holt (1957) and Winters (1960) non-seasonal exponential smoothing forecast method uses two smoothing coefficients, and incorporates an explicit linear trend in the forecast. Tourist arrival series from the origin country are smoothed in a recursive manner by weighting the current level and the previous smoothed estimate adjusted for the trend, as follows:

$$\tilde{A}_t = \alpha A_t + (1-\alpha)(\tilde{A}_{t-1} + T_{t-1}). \quad (4a)$$

The trend values are smoothed separately with a different smoothing coefficient, β. A recursive relation for the trend estimates is expressed as the weighted difference of two smoothed tourist arrivals series (which is equivalent to the slope of the smoothed series) and the previous trend:

$$T_t = \beta(\tilde{A}_t - \tilde{A}_{t-1}) + (1-\beta)T_{t-1}. \quad (4b)$$

The i-period-ahead forecast of tourist arrivals is given as follows:

$$F_{t+i} = \tilde{A}_t + iT_t, \qquad i = 1, 2, \ldots \quad (5)$$

The exponential smoothing forecast techniques are not appropriate for time series containing seasonality, as in the case of international tourism data. Quarterly international tourist arrivals to Australia from Hong Kong show pronounced seasonality. The Holt-Winters method can be extended to accommodate additive seasonality if the magnitude of the seasonal effects do not change with the series, or multiplicative seasonality if the amplitude of the seasonal pattern changes over time. These models involve three smoothing equations, with the forecast equation consisting of a constant, a linear trend and a seasonal factor:

$$F_{t+i} = \tilde{A}_t + iT_t + S_{t+i-j}, \quad i = 1,2, \ldots \quad (6)$$

where

$$\tilde{A}_t = \alpha(A_t - S_{t-j}) + (1-\alpha)(\tilde{A}_{t-1} + T_{t-1}), \quad 0<\alpha<1 \quad (7a)$$

$$T_t = \beta(\tilde{A}_t - \tilde{A}_{t-1}) + (1-\beta)T_{t-1}, \quad 0<\beta<1 \quad (7b)$$

and

$$S_t = \gamma(A_t - \tilde{A}_t) + (1-\gamma)S_{t-j}, \quad 0<\gamma<1 \quad (7c)$$

where α, β and γ are smoothing coefficients. Equation (7a) is similar to (4a) except that the most recent seasonally adjusted observation (obtained by subtracting the most recent seasonal observation from current tourist arrivals, $A_t - S_{t-j}$, and j = 4 for quarterly data) is used to obtain the smoothed tourist arrivals series. The estimate of the current additive seasonal factor, I_t, is obtained from equation (7c) for s periods per year, with j = 4. For the multiplicative seasonal model, equations (7a) and (7c) are rewritten as:

$$\tilde{A}_t = \alpha(A_t / S_{t-j}) + (1-\alpha)(\tilde{A}_{t-1} + T_{t-1})$$

and

$$S_t = \gamma(A_t / \tilde{A}_t) + (1-\gamma)S_{t-j}$$

where S_t is a multiplicative seasonal factor in the presence of a linear trend.

When an appropriate autoregressive integrated moving average (ARIMA) model has been fitted to a time series, the i-period-ahead forecast of tourist arrivals is given by:

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

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 after differencing (if necessary) are examined for stationarity. After empirical examination, the most appropriate model for tourist arrivals from Hong Kong is determined as ARIMA(3,1,1) (with absolute t-ratios in parentheses):

$$(1 + 1.32L + 0.82L^2 + 0.52L^3)(1 - L)\log HK_t$$
$$(10.7) \quad (4.20) \quad (4.15)$$
$$= 0.04 + \hat{\varepsilon}_{HKt} - 0.97\hat{\varepsilon}_{HK(t-1)}$$
$$(3.25) \qquad (38.4)$$

AIC = -3.65, SBC = -3.47.

Since the specific ARIMA model that adequately describes tourist arrivals from Hong Kong is as above, the fitted model used for ex post forecasting of tourist arrivals is given as follows:

$$F_{\log HK(t+i)} = 0.04 - 0.32\log HK_{t+i-1} + 0.50\log HK_{t+i-2} + 0.30\log HK_{t+i-3} + 0.52\log HK_{t+i-4} + \hat{\varepsilon}_{HK(t+i)} - 0.97\hat{\varepsilon}_{HK(t+i-1)}, \quad i = 1, 2, \ldots$$

The most appropriate multiplicative seasonal ARIMA model selected for the Hong Kong tourist arrivals series, namely ARIMA$(1,1,4)(0,1,0)_4$, can be used to forecast future observations, which are given as follows:

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

AIC = -3.98, SBC = -3.91.

3. Magnitude of Forecast Errors

The logarithm of quarterly tourist arrivals from Hong Kong is used to capture the multiplicative effects in the levels of the variables. With the final observation being tourist arrivals in the fourth quarter of 1989, Table 1 presents the RMSE one-quarter-ahead forecast accuracy measure of the single, double, Holt-Winters Nonseasonal, Holt-Winters Additive Seasonal, and Holt-Winters Multiplicative Seasonal exponential smoothing methods, and ARIMA models. The best fitting ARIMA and seasonal ARIMA models outperform all the other models, followed by the Holt-Winters Multiplicative Seasonal exponential smoothing method.

Using the best fitting ARIMA model for post-sample forecasting, the sample correlation of the actual and the corresponding forecasts is computed as a goodness-of-fit measure to determine how well the forecasts fit the actual values. This can provide useful information regarding how well the model forecasts. The correlation coefficient for the ARIMA model is 0.68. Even though the ARIMA model outperforms the other models in forecasting tourist arrivals from Hong Kong, only 68% of the variation in the tourist arrivals forecast is associated with variations in actual tourist arrivals between 1990(1) and 1996(4). The Lagrange multiplier test for serial correlation, LM(SC), shows that the null hypothesis of no serial correlation is not rejected.

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 in 1989-90. These one-off events might have distorted the estimation, testing and analysis of the underlying process, which is critical in forecasting. In order to analyse the impact of these events, intervention analysis with deterministic dummies allows the effects of these exogenous shocks to be represented by an ARIMA model, as follows (with absolute t-ratios in parentheses):

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

where
D1 = dummy variable equal to one for the 1979 Oil Price Crisis;
D2 = dummy variable equal to one for the 1988 Bicentennial Celebration;
D3 = dummy variable equal to one 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). Table 2 shows that none of the intervention variables used is significant at the 5% level, while the constant term, and the autoregressive and moving average coefficients remain significant.

4. Conclusion

This paper examined univariate time series forecasting methods based on current and past tourist arrivals in logarithms from Hong Kong to Australia, and computed various accuracy measures. Overall, 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 paper has concentrated on obtaining forecasts from time series methods of tourist arrivals from the own current and past values. Current and previous values of other related factors which could provide additional useful information, such as economic variables, will be incorporated into forecasts of tourist arrivals in future research.

Acknowledgements

The first author wishes to acknowledge the financial support of the Small and Medium Enterprise Research Centre and the School of Finance and Business Economics at Edith Cowan University. The second author wishes to acknowledge the financial support of the Australian Research Council and the Japan Society for the Promotion of Science.

References

Australian Bureau of Statistics, 1975-1996, Various Issues of Short-Term Movement: Overseas Visitors Arriving, Canberra.

Brown, R.G. (1963) *Smoothing, Forecasting and Prediction of Discrete Time Series*, Englewood Cliffs, NJ: Prentice Hall.

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

Granger, C.W.J. and Newbold, P. (1977) *Forecasting Economic Time Series*, New York: Academic Press.

Holt, C.C. (1957) Forecasting Seasonal and Trends by Exponentially Weighted Averages, Office of Naval Research, Memorandum 52.

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

Newbold, P. (1983) ARIMA Model Building and the Time Series Analysis Approach to Forecasting, *Journal of Forecasting*, 2, 23-35.

Winters, P.R. (1960) Forecasting Sales by Exponentially Weighted Moving Averages, *Management Science*, 6, 324-342.

Table 1

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

Model	RMSE
Single Exponential	0.24 (7)
Double Exponential	0.22 (6)
Holt-Winters	0.215 (5)
Holt-Winters Additive	0.13 (4)
Holt-Winters Multiplicative	0.12 (3)
ARIMA	0.016 (1)
Seasonal ARIMA	0.042 (2)

Note: Figures in parentheses denote rankings.

Table 2

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

Intervention Variables	Impact Effects
Oil Crisis 1979	0.07 (1.48)
Bicentennial Celebration 1988	-0.04 (-0.92)
Air Pilots Dispute 1989	0.11 (1.19)

Note: t-statistics are given in parentheses.