ABSTRACT

Japan is one of the most important international tourist source markets for Australia. As tourism forecasting plays a critical role in tourism planning, it is imperative for policymakers in both the public and private tourism sectors in Australia to examine tourist arrival patterns from Japan, and to generate forecasts for this inbound market. The objectives of this paper are to determine suitable forecasting models for Japanese tourist arrivals to Australia and to forecast tourist flows using the ‘best’ estimated models. In this study, the Box-Jenkins models will be compared with other seasonal time series models.

1 INTRODUCTION

Forecasting tourism demand is important for tourism planning at all levels in the tourism industry from the government to a single tourist business. The value of forecasting lies in its ability to reduce the loss caused by disparities between demand and supply. In order to provide satisfactory services to tourists, destinations need to acquire reliable forecast of future demand for transportation, accommodation, service personnel, tourist retailers and other facilities.

Japan is one of the most important international tourist source markets for Australia. As tourism forecasting plays a critical role in tourism planning, it is imperative for policymakers in both the public and private tourism sectors in Australia to examine tourist arrival patterns from Japan, and to generate forecasts for this inbound market. Lim and McAleer (2001) examine international tourism demand by Japan for Australia, as measured by tourist arrivals from 1976 to 1999 using Box-Jenkins models. As an extension of their study, the objectives of this paper are to determine suitable forecasting models for Japanese tourist arrivals to Australia and to forecast tourist flows using the ‘best’ estimated models. In this study, the Box-Jenkins models will be compared with other seasonal time series models.

2 JAPANESE OUTBOUND TOURISM TO AUSTRALIA

Since the first relaxation of overseas travel restriction, the number of Japanese outbound tourists has grown rapidly to more than 16 million in 2001 (JTB, 2003, January 9). Besides, Japanese tourists have acquired a reputation for being lavish spenders. In 2001, Japan was the fourth largest spender in terms of international tourism expenditures (US$26.5 billion) and Japanese international travelers made up 5.7 percent of the total international tourism market (World Tourism Organisation, 2002). The rapid growth of the Japanese outbound tourism was mainly attributed to the country’s strong economy. In addition, lifestyle change, strong Japanese Yen in the 1980s and early 1990s, and government encouragement of overseas travel were some of the most prominent factors contributing to the boom of Japanese outbound tourism. The Japanese outbound travel market has been in recession since 1997 due to the slowdown of the Japanese economy. Weaker Japanese Yen, fall in consumer confidence, and concerns over job security have caused Japanese tourists to delay their overseas holidays.

Australia has long been one of the most popular overseas destinations for Japanese travelers. In 2000, Australia which ranked seventh, had about 700,000 tourists from Japan (The Judicial System and Research Department, Ministry of Justice, Japan, 2001, March). Japan which was in Australia’s list of top ten international tourist source markets, rose from fifth place in 1974 to rank at number 1 in 1990. It had remained in the first place for almost a decade, making it Australia’s single largest source of international tourists in the 1990s (Buchanan, 1999). However, this market has been in recession since September 1997 and tourist arrivals from Japan to Australia started to decline.
from a peak of over 800,000 in 1997. In 1999, tourists from New Zealand to Australia superseded tourist arrivals from Japan and became Australia’s largest tourist source country. In the meantime, Hawaii overtook Australia to be the most preferred destination that Japanese tourists would like to visit. Moreover, Australia is facing stronger competition from China, the U.S. mainland, Guam and Hong Kong (March, 2003). The sharp declines in the Japanese outbound market and Japanese tourist arrivals to Australia since 1997 were caused mainly by the Asian economic and financial crises. The continuing economic slowdown in Japan has inevitably affected the country’s demand for international travel.

Figure 1 is the time plot of Japanese tourist arrivals to Australia. The key feature is its noticeable upward trend before 1997 and downward trend thereafter. As mentioned above, the turning point of the data series corresponds to the Asian economic and financial crises in 1997. The time series also display seasonal patterns. More Japanese tourists travelled to Australia in the first and fourth quarters than in the second and third quarters, especially in the early 1990s onward. Most of them are leisure tourists on package tours. The most popular State visited by Japanese tourists in 2003 was Queensland, followed by New South Wales (Australian Tourist Commission, 2004).

![Figure 1: Quarterly Japanese tourist arrivals to Australia, 1976(1)-2000(2)](image)

### 3 FORECASTING ACCURACY MEASUREMENTS

In this study, the mean absolute percentage error (MAPE), mean percentage error (MPE), root mean squared error (RMSE), and Theil’s $U$ are used to measure and interpret the forecast errors, as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100$$  \hspace{1cm} (1)

$$\text{MPE} = \frac{1}{n} \sum_{t=1}^{n} \left( \frac{Y_t - \hat{Y}_t}{Y_t} \right) \times 100$$  \hspace{1cm} (2)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2}$$  \hspace{1cm} (3)

$$\text{Theil’s } U = \frac{\sum_{t=1}^{n} (Y_t - \hat{Y}_t)^2}{\left( \sum_{t=1}^{n} (Y_t) - \frac{1}{n} \sum_{t=1}^{n} \hat{Y}_t \right)^2} \div \left( \sum_{t=1}^{n} (Y_t)^2 - \frac{1}{n} \left( \sum_{t=1}^{n} Y_t \right)^2 \right)$$  \hspace{1cm} (4)

RMSE is the square root of the average of all squared errors. Although this measurement ignores any over- and under-estimation, RMSE does not allow for the comparison across different time series and for different time intervals. However, MAPE and MPE, which are percentage error measurements, allow for such comparison. MAPE is particularly useful when the units of measurement of $Y$ are relatively large. MPE takes into account whether a forecasting method is biased (that is, the forecasts are consistently lower or higher than the actual observations). The Theil’s $U$ statistic is a relative measurement based on a comparison of the predicted change with the observed change. The value of $U$ lies between 0 and 1. If $U$ equals to 0, there is a perfect fit, whereas $U$ equals to 1 implies that the forecasting ability of a particular model is zero.

### 4 FORECASTING MODELS

Given that the quarterly tourist arrivals from Japan to Australia exhibit trend and seasonal patterns (see Figure 1), the forecasting techniques used in the study incorporate both trend and seasonal components in the data series. Four time series forecasting models which accommodate seasonality are considered in this study. They include the Naïve Trend and Seasonal, Time Series Decomposition, Winter’s Exponential, and SARIMA models.

#### 4.1 Naïve Trend and Seasonal model

The Naïve Trend and Seasonal model incorporates both seasonal and trend variations. The model is given in equa-
Wang and Lim

4.2 Winter’s Exponential models

Winter’s (1960) three-parameter exponential smoothing method allows for evolving linear trend and seasonality. Different smoothing constants are used to directly smooth the level, slope, and seasonal effect in a series, as shown in equations (6), (7), and (8), respectively, as follows:

\[ L_t = \alpha \frac{Y_t}{S_t} + (1 - \alpha)(L_{t-1} + T_{t-1}) \]  
\[ T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \]  
\[ S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s} \]

where \( L_t \) = current level estimate, \( \alpha \) = smoothing constant for the current level estimate, \( Y_t \) = value in period \( t \), \( \beta \) = smoothing constant for trend estimate, \( T_t \) = trend estimate, \( \gamma \) = smoothing constant for seasonality estimate, \( S_t \) = seasonality estimate, \( s \) = length of seasonality.

Equation (9) is used to estimate forecast errors and to generate ex ante forecasts:

\[ \hat{Y}_{t+p} = (L_t + pT_t)S_{t+p} \]  

where \( p \) = periods to be forecast into the future.

4.3 Time Series Decomposition method

This approach attempts to decompose a time series into trend (\( T \)), seasonal (\( S \)), cycle (\( C \)), and irregular components (\( I \)). The purpose of this approach is to distinguish these four components, develop forecasts for each component and then recombine them to produce forecasts (Frechtling, 2001). Equation (10) is the multiplicative form of the model, which assumes a multiplicative relationship among the components.

\[ Y_t = T_t \times S_t \times C_t \times I_t \]  

4.4 Box-Jenkins ARIMA and SARIMA models

The ARIMA model, also known as the Box-Jenkins methodology, is the most popular advanced extrapolative method to date. In this analysis, a series is transformed into a stationary series if necessary before it is identified, estimated, diagnosed, and forecasted. One of the advantages of this method is that various ARIMA models can represent a wide range of characteristics of time series that occur in practice. Another important feature of the method is the readily available formal model checking procedure, which enable diagnostic checkings to be carried out to examine the validity or adequacy of the estimated models (Hanke et al., 2001).

ARIMA(\( p,d,q \)) stands for autoregressive(AR) integrated(I) moving average(MA), and (\( p,d,q \)) represents the order of autoregression, differencing and moving average, respectively. The ARIMA model can be expressed as follows:

\[ Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \ldots + \Phi_p Y_{t-p} + \varepsilon_t \]

\[ -\theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \]  

where \( Y_t \) is the number of tourist arrivals from a particular origin to a destination at time \( t \); \( Y_{t-1}, Y_{t-2}, \ldots Y_{t-p} \) are the lagged tourist arrivals at time lags \( t-1, t-2, \ldots t-p \); \( \Phi_0, \Phi_1, \ldots \Phi_p \), and \( \theta_1, \theta_2, \ldots \theta_q \) are the coefficients to be estimated; \( \varepsilon_t \) is the error term at time \( t \), which represents the effect of variables unexplained by the model; \( \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots \varepsilon_{t-q} \) are errors in the previous time periods.

Seasonal ARIMA or SARIMA, is an extension of ARIMA where seasonality is accommodated by seasonal differencing. SARIMA(\( p,d,q)(P,D,Q \)) is applied to data with an annual seasonal pattern. \( P, D, \) and \( Q \) in the seasonal part of the model represent the order of autoregression, differencing and moving average, respectively. As quarterly data are used, the seasonal difference for period \( S = 4 \) is defined as:

\[ \Delta_4 Y_t = Y_t - Y_{t-4} \]

5 UNIT ROOT TESTS

Before fitting ARIMA models to the tourist arrival series, logarithmic transformation is applied to the series to capture the multiplicative effect in the level of tourist arrivals. The graphical representation of the natural logarithmic
transformation of tourist arrivals from Japan to Australia is given in Figure 2.

![Figure 2: Logarithm of quarterly Japanese tourist arrivals to Australia, 1976(1)-2000(2)](image)

A stationary time series is one with constant mean and variance over time (Hill et al., 2001). The logarithmic tourist arrival series in Figure 2 appears to be non-stationary. Visual analysis of the correlograms of the log series also suggests that the logarithmic tourist arrivals are non-stationary, as the autocorrelations are typically large for the first several time lags and gradually drop towards zero.

The Augmented Dickey-Fuller (ADF) tests for unit root are used for the logarithmic tourist arrivals to Australia. The ADF test allows a parametric correction for higher-order serial correlation and is expressed as follows:

$$\Delta Y_t = \alpha_0 + \beta T + \gamma Y_{t-1} + \sum_{i=1}^{k} \alpha_i \Delta Y_{t-i} + \nu_t$$

(13)

where $\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2}), \Delta Y_{t-2} = (Y_{t-2} - Y_{t-3}), \ldots \ldots T = \text{deterministic trend.}$

Under the null hypothesis of a unit root, that is $\gamma = 0$, the ADF statistic does not follow the conventional $t$-distribution. If the absolute ADF test statistic is less than the absolute critical value, the series has a unit root and is said to be non-stationary (Hill et al., 2001). In order to determine the lag length $k$, an initial lag length of 4 is used in the ADF regression, and the fourth lag is tested for significance using the standard $t$-test at the 5% level. If the fourth lag is insignificant, the lag length is subsequently reduced until a significant lag length is obtained. The ADF test statistic for the logarithmic Japanese tourist arrivals to Australia is greater than the critical value and the null hypothesis of a unit root is not rejected at the 5% level of significance for the series. However, the ADF test statistics for the first differenced logarithmic Japanese tourist arrivals is (marginally) smaller than the critical value and the null hypothesis of a non-stationary series is rejected (see Table 1).

<table>
<thead>
<tr>
<th>Series</th>
<th>ADF lag length</th>
<th>ADF statistic</th>
<th>ADF critical statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log (Australia)</td>
<td>4</td>
<td>-0.17</td>
<td>-3.46</td>
</tr>
<tr>
<td>DLog (Australia)</td>
<td>4</td>
<td>-2.96</td>
<td>-2.90</td>
</tr>
</tbody>
</table>

Note: A deterministic trend is included in the ADF auxiliary regression for unit root test for log levels. The critical value is at the 5% significance level.

6 EMPIRICAL FORECASTING ANALYSIS

Quarterly data on Japanese tourist arrivals to Australia for 1976(1) to 2000(2) are obtained from the Australia Bureau of Statistics (Australia Bureau of Statistics, 1976-2000). This study also uses ex post forecast whereby the last 8 quarters of the series are reserved as ex post period, whereas the rest of the series are used as sample period for model estimation. The EViews 4 (2000) computer software package is used to estimate the exponential smoothing and SARIMA models.

The smoothing constants in Winter’s Exponential models are estimated by minimizing the sum of squared error (SSE) and the root mean squared error (RMSE) of the forecast. The mean of the initial observations is used to start the iterative calculations by the software. Using Time Series Decomposition method, a linear trend equation is first estimated using least square method, which minimizes the SSE. After removing the trend component from the series, the seasonal index is estimated by taking the median values of the particular quarters, which is adjusted by the seasonal multiplier. An estimate of the cyclical effect ($C$) is computed using a three-quarter moving average of the CI values. In ex post and ex ante forecasting, the cyclical
component is also calculated by three-quarter moving average, and the irregular effect is set at 1.

Various SARIMA models are fitted to the logarithms of tourist arrivals series from Japan to Australia, using regular and seasonal differencing at lag 4. Two criteria are used in the initial selection of the appropriate model. First, the AR and MA coefficients must be statistically significant at the 5% level. Second, the absence of serial correlation in the residuals, using the Lagrange multiplier test for serial correlation. The selected models are then compared according to the Akaike Information Criterion (AIC), the Schwarz Bayesian Criterion (SBC) and the residual correlogram. The model selected should have the smallest AIC and SBC values and random residuals.

By using different orders for \( p, q, P, \) and \( Q \) from 0 to 4, six models have been identified with significant parameters to explain Japanese tourist arrivals to Australia. The LM tests for all six models fail to reject the null hypothesis of no serial correlation. Residual checking using \( p \)-value associated with the \( Q \)-statistic also shows that the residuals of the six models are random. Based on the AIC criterion, SARIMA\((2,1,3)(0,1,4)\) is the best model while SARIMA\((0,1,3)(0,1,2)\) is the best model using the SBC criterion. However, the latter is not useful for forecasting tourist arrivals from Japan to Australia because it does not contain the AR and SAR components of the SARIMA model (see Table 2). SARIMA\((2,1,3)(0,1,4)\) and SARIMA\((0,1,3)(0,1,2)\) are given in equations 15 and 16, respectively:

\[
\Delta_i Y^d_t = -0.0043 + 0.7722 Y_{t-2} + \varepsilon_t + 0.3060 \varepsilon_{t-1} + 0.9452 \varepsilon_{t-2} - 0.2554 \varepsilon_{t-3} + 0.4091 \varepsilon_{t-16}
\]

(14)

\[
\Delta_i Y^d_t = -0.0022 + \varepsilon_t + 0.3190 \varepsilon_{t-1} + 0.9033 \varepsilon_{t-2} - 0.2319 \varepsilon_{t-3} - 0.6698 \varepsilon_{t-8}
\]

(15)

where \( \Delta_i Y^d_t \) is the 1st differenced and seasonally differenced logarithm of the quarterly Japanese tourist arrivals to Australia.

Table 2: Estimated SARIMA model for Japanese tourist arrivals to Australia

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-statistic</th>
<th>AIC/SBC</th>
<th>LM(SC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.0043</td>
<td>-2.24</td>
<td>AIC= -1.47</td>
<td>F=0.85</td>
</tr>
<tr>
<td>AR(2)</td>
<td>0.7722</td>
<td>6.86</td>
<td>BSC= -1.29</td>
<td>P=0.43</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.3060</td>
<td>-2.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA(2)</td>
<td>-0.9452</td>
<td>-11.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA(3)</td>
<td>0.2554</td>
<td>1.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMA(4)</td>
<td>-0.4091</td>
<td>-3.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0022</td>
<td>-2.14</td>
<td>AIC= -1.46</td>
<td>F= 1.25</td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.3190</td>
<td>-2.86</td>
<td>P=0.29</td>
<td></td>
</tr>
<tr>
<td>MA(2)</td>
<td>-0.9022</td>
<td>-18.34</td>
<td>BSC= -1.32</td>
<td></td>
</tr>
<tr>
<td>MA(3)</td>
<td>0.2319</td>
<td>2.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMA(2)</td>
<td>0.6698</td>
<td>7.70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7 FORECASTING PERFORMANCE

For each model used in this study, four-quarter and eight-quarter-ahead \textit{ex post} forecasts are calculated and compared with the actual values of the series. MAPE, MPE, RMSE, and Theil’s \( U \) are used to assess forecast accuracy. The models which have the lowest forecast errors are identified as the “best” models. The empirical results of the forecasting performance of the various models are given in Table 3. The ranking of each model is also provided in the column next to the value of the accuracy measurements. It is clear from the table that Winter’s Exponential and SARIMA models are the ‘best’ models for the four-quarter and eight-quarter-ahead forecasts, respectively, for Japanese tourists to Australia.
8 EX ANTE FORECASTS

The “best” models for each series obtained above are used to generate four-quarter and eight-quarter-ahead forecasts, respectively. Winter’s Exponential and SARIMA models are re-estimated using all the observations in the series. The re-estimated coefficients of the parameters in the SARIMA (2,1,3)(0,1,4) model for Japanese tourist arrivals to Australia are given in Table 4 and the model is expressed as follows:

\[\Delta_4 Y_t = -0.0039 + 0.7764 Y_{t-2} + \varepsilon_t + 0.3104 \varepsilon_{t-1} + 0.9468 \varepsilon_{t-2} - 0.2775 \varepsilon_{t-3} + 0.4193 \varepsilon_{t-16}\]  

(17)

Although the constant term in the re-estimated SARIMA model is not significant at the 5% level, it is almost significant at the 10% level. The re-estimated SARIMA model has no serial correlation and satisfies the diagnostic check. The re-estimated coefficients \(\alpha, \beta, \text{ and } \gamma\) in the Winter’s exponential models for the series of Japanese tourist arrivals to Australia are 0.4001, 0.1599, and 0.8902, respectively. The ex ante forecasts for Japanese tourist arrivals to Australia are given in Table 5.
The ex ante forecasts suggest that the flow of Japanese tourists to Australia for each respective quarter will continue to fall, following the decline in Japanese tourist arrivals since 1998. As indicated by both Winter’s exponential and SARIMA models, seasonal fluctuations will remain to be an important feature in Japanese tourist arrivals, with some variations between two consecutive quarters exceeding 10 percent.

9 CONCLUSION

As planning is a crucial aspect of tourism for policymakers, the importance of forecasting tourism demand cannot be overlooked. Given the fact there are very few studies on Japanese outbound tourism to Australia, a time series modelling and forecasting of Japanese outbound tourism to Australia is presented in this paper.

Various time series forecasting models accommodating seasonality have been used to explain Japanese quarterly tourist arrival patterns to Australia for the period 1976(1) to 2000(2). Before estimating the SARIMA models, the tourist arrivals in log levels have been transformed into stationary series. The ADF tests for unit roots are used to check the stationarity of the log series before and after ordinary and seasonal differencing for tourist arrivals to Australia. The AIC and SBC are used to obtain the best fitting SARIMA models for the series. Additionally, the LM test for serial correlation and residual correlogram are used to verify the adequacy of the estimated SARIMA models.

The performances of the various models are evaluated by analysing their ex post forecast accuracy using MAPE, MPE, RMSE, and Thiel’s $U$ accuracy measurements. Winter’s Exponential and SARIMA models are the best models for forecasting four-quarter and eight-quarter-ahead, respectively for Japanese tourist arrivals to Australia.

The best models for each forecasting horizon are subsequently used to generate ex ante forecasts for tourist arrivals from Japan to Australia. As the forecasts provide some important information about the changing patterns in the Japanese market, it is useful for both the public and private sectors in Australia to incorporate this in their policy decision making to build a sustainable tourism industry.

Apart from identifying the major factors causing the possible future decline in Japanese tourist arrivals, tourism authorities in Australia also need to identify new tourist segments in Japan that have growth potential for Australia. Developing effective marketing strategy and appropriate tourism products are vital for rejuvenating the Japanese market.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support of Griffith University.

REFERENCES


AUTHOR BIOGRAPHIES

Ying Wang is a first year PhD student at Griffith University, Australia. She is a recipient of the Griffith Business School Postgraduate Research Scholarship. Her research interest is in modelling international tourism demand by
China for Australia. Her email address is <ying.wang@griffith.edu.au>

Dr Christine Lim is an Associate Professor at Griffith University, Australia. She obtained her PhD from the University of Western Australia. Her research interests are in applied econometrics, tourism modelling and management.

Christine is a Resource Editor of Annals of Tourism Research, an Associate Editor of the Journal of Hospitality and Tourism, and an Executive Committee member of the Modelling and Simulation Society of Australia and New Zealand. Her email address is <christine.lim@griffith.edu.au>