Malaysia is well-known for her ‘blue’ and ‘green’ tourism attractions. Malaysian government launched several tourism programs to encourage and attract international tourist arrivals into Malaysia. This study therefore attempts to forecast the tourism demand for Malaysia from ASEAN countries. The literature on forecasting tourism demand is huge comprising various types of empirical analysis. Some of the researchers applied cross-sectional data, but most of the tourism demand forecasting used pure time-series analytical models. One of the important time-series modelling used in tourism forecasting is ARIMA modelling. This study employs quarterly time series data of ASEAN tourist arrivals to Malaysia for the period from 1995:Q1 to 2009:Q4 to forecast future tourism demand for Malaysia. The forecasting performance is based on seasonal ARIMA model. The findings of this study revealed that seasonality model does not offer any valuable insights or provide reliable forecasts on tourism demand for Malaysia by ASEAN countries. This scenario occurs because of the fact that ‘Malaysia is Truly Asia’.

Keywords: one-period-ahead forecasting, SARIMA, ASEAN

JEL Classification: L83, M1, O1

INTRODUCTION

Malaysia is one of Asia’s most popular tourist destinations. Many studies were undertaken to assess international tourist perceptions of
Malaysia as a tourist destination. Lower real exchange rate and political stability made Malaysia as an affordable tourist destination. Malaysia is also well known for its ‘green tourism’ with attractive tropical environment; ‘blue tourism’ with beautiful beaches and islands; historical and culinary attraction; diverse culture; diverse ethnic food (see for example Josiam, Sohail & Monteiro, 2007); world-class hotels and resorts; and excellent shopping places. East Asia, which include Japan, Korea, China and Taiwan, including ASEAN (Association of Southeast Asian Nations) countries are Malaysia’s major tourist market, with a market share of 75% of the overall international tourist arrivals to Malaysia in 2010. Singapore, Indonesia and Thailand are three major tourism source markets for Malaysia, with a total share of 60% from the total international tourist arrivals to Malaysia in 2010. This phenomena concurs with ‘Malaysia is Truly Asia’ tagline promoted by Tourism Malaysia. In spite of the phenomenal growth in inbound tourism from ASEAN region, little research has been undertaken on these markets to evaluate their contributions to Malaysia’s inbound tourism industry. The international tourism can be estimated in terms of the number of international tourist arrivals. Meanwhile, domestic tourism generally is not considered for most of empirical analysis purposes. The World Tourism Organization predicted that there will be 1.6 billion international tourist arrivals worldwide by 2020 and that these tourists are also expected to spend over two trillion US dollars (WTO Report, 2010). Thus, the discussion of this paper is organised as follows. The second section of this paper will discuss the selected literature review. The third section focuses on methodology used. The fourth section touches on empirical findings and the final section concludes.

LITERATURE REVIEW

Along with the phenomenal growth in demand for tourism worldwide over the past two decades, there is also a growing interest in tourism research. The literature on modelling and forecasting tourism demand is numerous with various type of empirical analysis. Some of the researchers applied cross-sectional data, but most of tourism demand forecasting used pure time-series analytical models. One of the important time-series modelling used in forecasting tourism demand is ARIMA modelling, which is specified based on standard Box-Jenkins method, a famous modelling approach in forecasting demand. Many studies have applied this methodology, such as Chu (2008a), Lee, Song and Mjelde (2008), Coshall (2009), Wong et al. (2007), Akal (2004), Preez and Witt
(2003); and Kulendran and Witt (2001). The ARIMA model is proven to be reliable in modelling and tourism demand forecasting with monthly and quarterly time-series. Wong et al. (2007) used four types of models, namely seasonal auto-regressive integrated moving average model (SARIMA), auto-regressive distributed lag model (ADLM), error correction model (ECM) and vector-autoregressive model (VAR) to forecast tourism demand for Hong Kong by residents from ten major origin countries. The empirical results of the study shows that forecast combinations do not always outperform the best single forecasts which have been used frequently in previous studies. Therefore, combination of empirical models can reduce the risk of forecasting failure in practice. For example, Coshall (2009) used univariate time series approach, combining the ARIMA, volatility and smoothing models to forecast United Kingdom’s demand for international tourism. Generally, from this study we can conclude that the ARIMA volatility models tend to overestimate demand, and the smoothing models are inclined to underestimate the number of future tourist arrivals.

Chu (2008a) modified ARIMA modelling to fractionally integrated autoregressive moving average (ARFIMA) in forecasting tourism demand. This ARFIMA model is ARMA based methods. Three types of univariate models were applied in the study with some modification in ARMA model to become ARAR and ARFIMA model. The main purpose of this study is to investigate the ARMA based models and its usefulness as a forecast generating mechanism for tourism demand for nine major tourist destinations in the Asia-Pacific region. This study is different from other tourism forecasting studies published earlier, because we can identify the ARMA based models behaviour and the difference between ARFIMA models with other ARMA based models. Again, Chu (2008b) studied the ARIMA based model using ARAR algorithm model in order to analyze and forecast tourism demand for Asia-Pacific region using monthly and quarterly data. This study reveals that the performance between forecasts generated by monthly and quarterly data is similar. Besides forecasting tourist arrivals, prediction of tourism revenue can also be done using empirical modelling. Akal (2004) used autoregressive integrated moving average cause-effect (ARMAX) modelling to forecast international tourism demand for Turkey. The ARMAX model is actually derived from the ARIMA approaches. The forecast estimations are an important benchmark for Turkey’s government to strengthen the tourism sector in order to transform it into a major contributing sector for economic development in the future.
Song, Wong and Chon (2003) introduced general-to-specific modelling approach to forecast international tourist arrivals from 16 major countries to Hong Kong for the period from 2001 to 2008. The specification of econometric model used is known as auto-regressive distributed lag model (ADLM). This study shows that, several important factors that determine the demand for Hong Kong tourism include the cost of tourism in Hong Kong, the economic scenario in the origin countries, the costs of tourism in the competing destinations and the ‘word of mouth’ effect. Again, ADLM measurement of tourism forecasting is suitable for multivariate modelling and by using this method we are able to determine various factors that cause tourist arrivals in the future. Greenidge (2001) used structural time-series modelling (STM) to evaluate tourism demand forecasting in Barbados. STM modelling has its own strength, where time-varying components can be incorporated in the regression equation to capture the movement of tourist arrivals using explanatory variables. Besides using basic structural modelling (BSM), STM model is also able to include general structural modelling (GSM) with seasonal effect. Loganathan, Ibrahim and Harun (2008) on the other hand, used standard Johansen-Juselius co-integration method to study the importance of the tourism industry in enhancing trade performance and economic development in Malaysia. Thus, these studies provide valuable insights into the stylized facts of tourism behaviour and provide reliable out-of-sample forecasts of tourism demand.

Athanasopoulos and Hyndman (2008) modelled Australian domestic tourism demand using regression model, exponential smoothing via innovations state space model and innovations state space model with exogenous variables. Cross-sectional data was applied in this study, and the data was collected using computer-assisted telephone interviews from 120,000 Australians aged 15 years and above. All the models used in the study highlighted the impact of world events on Australian domestic tourism. Such events are the increase in business travel immediately after the Sydney Olympic in the year 2000 and the significant increase in visiting relatives and friends after the 2002 Bali bombings. One interesting aspect that can be found from this study is that, all three statistical models used in this study outperform the Tourism Forecast Committee (TFC) results for short-term demand of Australian domestic tourism. Meanwhile, the long-term forecasts results from this study also indicate that the TFC forecasts may be too optimistic. Finally, from the forecasts outputs, this study found that the Australian domestic tourism is on the declining stage. Psillakis, Panagopoulos and Kanellopoulos (2009) forecasted tourism demand in accommodation industry in Greece using
implemented forecasting model and Box-Jenkins approach. Both approaches used in this study have its own strength, where the forecasted values indicate expected value of the predictor of a future time series value. Botti et al. (2007) also used a simple econometric model to study the demand for tourism in France, and concluded that tourism demand in France is dependent on available income and relative prices.

Unlike most of the tourism forecasting studies discussed earlier, forecasting expo demand involves both qualitative technique and quantitative forecasting models (Lee, Song & Mjelde, 2008). The main reason to use both techniques is because of the limitation of data availability. Combining quantitative technique with willingness-to-visit (WTV) surveys, the number of visitors expected to visit international tourism expo to be held in Korea in 2012 is predicted. Preez and Witt (2003) compared two types of methods to analyze tourism forecasting. In their study, univariate and multivariate modelling were used separately to forecast tourism demand from four European countries to Seychelles. The findings of the study revealed that the univariate forecasting models outperform compare to multivariate models. The empirical result shows that ARIMA estimation exhibits better forecasting performance than univariate and multivariate state space modelling. According to Kim and Wong (2006), the volatility in tourism demand data can be influenced by the effects of new shocks such as economic crises, natural disaster or war. In tourism literature, modelling the volatility in tourism demand is important because it can capture the occurrence of unexpected events. Volatility of tourism demand is modelled using conditional volatility models, and the models that appears in tourism literature are univariate generalized autoregressive conditional heterokedasticity (GARCH), univariate asymmetric GARCH, vector autoregressive moving average GARCH (VARMA-GARCH); and VARMA asymmetry GARCH (VARMA-AGARCH) models (Chan, Lim & McAleer (2005); Kim & Wong (2006); Shareef & McAleer (2005); Shareef & McAleer (2007).

In middle of 1990s, dynamic specification such as vector autoregressive model (VAR), error correction model (ECM) and autoregressive distributed lag model (ADLM) began to appear in the tourism literature. VAR model is able to accommodate various types of independent variables to determine tourist arrivals and from there we are able to forecast future tourist arrivals. Besides that, VAR model enables innovative use of impulse response analysis in tourism context, besides providing results on co-integrating analysis and forecasting. Song and Witt (2004) used VAR model to forecast international tourist flows to Macau for the period from 2003 to 2008. The forecast generated by VAR
models suggest that Macau will face increasing tourism demand by residents from mainland China. The ECM model had also been used to measure tourism forecasting but in the recent past, ECM model had been modified into vector error correction model (VECM) which can test and impose weak exogeneity restrictions. Bonham, Gangnes and Zhau (2008) used VECM technique to identify reasonable long-run equilibrium relationship and took into account Diebold-Mariano tests to demonstrate satisfactory forecasting performance for Hawaii. Beyond the forecasting approach used in terms of tourism demand, Akal (2010) investigated the economic implications of international tourism on Turkish economy. The increasing flows of international tourist according to the author will have many economic implications. In Turkey for instance, the dynamic flows of international tourists have increased accommodation facilities; create new employment in the tourism related industry; increased foreign direct investment in tourism related industry; and foreign currency has become major source of Turkish economy.

The main purpose of this study is to provide a much more comprehensive examination on tourism forecasting using seasonal ARIMA modelling. However, existing literature on forecasting international tourism demand for Malaysia so far had not adopted seasonal auto-regressive integrated moving average (SARIMA) modelling. Therefore, this paper intends to fill this gap.

DATA AND MODEL SPECIFICATION

Data for this study was collected from Malaysia tourism arrivals dataset provided by the Ministry of Tourism Malaysia. The time-series data used in this study is quarterly data and it covers a period from 1995:Q1 to 2009:Q4. This study focuses on the demand for tourism in Malaysia from ASEAN countries and forecast 4 quarters ahead, which is 2010:Q1 to 2010:Q4. In this study, we used seasonal ARIMA models to forecast one-period ahead of the series by applying Box-Jenkins approach. An ARIMA model is a generalization of an ARMA model. These models are fitted to time-series data either to better understand the data or to predict future points in the series (Chu, 2008a). The model is generally referred to as an ARIMA(p,d,q) model where p,d and q are integers greater than or equal to zero and refers to the order of the autoregressive, integrated and moving average aspects. In this study we applied Augmented Dickey-Fuller (ADF) and Phillip-Perron (PP) stationary tests to identify the level d in time series of international tourist
arrivals to Malaysia. The resulting univariate time series model can be written as follows:

\[ \Phi_p(L)y_t = \Phi_q(L)\epsilon_t, \ t= p+1, p+2, \ldots, n \]  
(1)

with;

\[ \Phi_p(L) = 1- \Phi_1L - \cdots - \Phi_pL^p \]
\[ \Phi_q(L) = 1- \Phi_1L - \cdots - \Phi_qL^q \]  
(2)

Where, the last notational conventional is chosen such that the model in (1) amounts to the following regression model:

\[ Y_t = \Phi_1y_{t-1} + \cdots + \Phi_py_{t-p} + \epsilon_t + \varphi_1\epsilon_{t-1} + \cdots + \varphi_q\epsilon_{t-q} \]  
(3)

Therefore, this model is called an autoregressive moving average model of order (p,q), or briefly ARMA(p,q). When the \( y_t \) series replaced by \( \Delta^d y_t \), we say that \( y_t \) is described by an autoregressive integrated moving average model of order (p,d,q), or briefly ARIMA(p,d,q). This can be expressed as Box-Jenkins approach. The autocorrelation functions (ACF) of a time series \( y_t \) can be define as \( \rho_k = \gamma_k / \gamma_0 \). Where \( \gamma_k \) is the k order of auto-covariance of \( y_t \), that is

\[ \gamma_k = E[(y_t-\mu)(y_{t-k}-\mu)], \ k = \ldots, -2, -1, 0, 1, 2, \ldots, n \]  
(4)

Given equation (3), it can be seen that autocorrelation holds when \( \rho_0=1, \rho_k=\rho_k \) and that \( -1<\rho_k<1 \). Therefore, ACF can be useful to characterize ARIMA time series models. For example, taking into account a simple white noise series \( \epsilon_t \) for which \( E(\epsilon_t)=0 \) and \( \rho_k=0 \) for all \( k\neq0 \), the AR(1) model is:

\[ y_t - \mu = \phi_1(y_{t-1}-\mu) + \epsilon_t, \ t = 1,2,3,\ldots,n \]  
(5)

Meanwhile, the ACF may not be particularly useful to identify whether an AR of specific order is a suitable model. In fact, the ACF is more useful in the case of MA(q) models. For the MA(1) process, it can be shown as follows with PACF at lag h:

\[ \alpha(h) = \phi_{hh} = (-\theta)^h / (1 + \theta^2 + \cdots + \theta^{2h}) \]  
(6)
Forecasting SARIMA processes is completely analogous to the forecasting of ARIMA processes. This can be expressed as ARIMA(p,d,q)4 for quarterly data. Where, ‘p’ indicates the order of autoregressive operator; ‘d’ are the differences; and ‘q’ are the orders of moving average operator of non-seasonal and seasonal components respectively. The first two steps in identifying SARIMA models for a data set are to find ‘d’ and to create the differenced observations:

\[ y_t = (1 - B)^d (1 - B^4)^D X_t \]  

(7)

In this study we used one-period-ahead forecasting using seasonal ARIMA modelling. In order to forecast SARIMA model, the mean absolute percentage error (MAPE) is a useful measure to compare the accuracy of forecasts between different forecasting models since it measures relative performance. If an error is divided by the corresponding observed value, we have a percentage error. In many empirical studies it appears that, models that tend to do best for within-sample data do not necessarily forecast better in out-of-the sample. There is no strict rule for that, but empirical experience suggests that it would be better to select several models based on the Akaike Information Criterion (AIC) and evaluate these on the forecasted data. The last evaluation can be based on root mean square error (RMSE). The RMSE can be expressed as follows:

\[ \text{RMSE} = (1 - m) \left\{ \sum_{h=1}^{m} (\hat{y}_{n+h} - y_{n+h})^2 \right\} \]  

(8)

Meanwhile, in previous literatures, mean absolute percentage error (MAPE) was used to determine suitable models. It worth mentioning, that MAPE is not very useful for very small observation (Fransese, 1998). The MAPE can be expressed as follows:

\[ \text{MAPE} = (1 - m) \left\{ \sum_{h=1}^{m} \left| (\hat{y}_{n+h} - y_{n+h}) / y_{n+h} \right| \right\} \]  

(9)

Therefore, ARIMA-SARIMA model selection in this study is based on AIC forecast evolution results, especially referring to the minimum value of RMSE and MAPE.
EMPIRICAL RESULTS

Table 1 summarizes the outcome of the ADF (Augmented Dickey-Fuller) and PP (Phillip-Perron) tests on ASEAN tourist arrivals to Malaysia. The null hypothesis tested is that the variable under investigation has a unit root against the alternative that it does not. The lag-length is chosen using the AIC after testing for first and higher order serial correlation in the residuals. In Table 1, the null hypothesis cannot be rejected by both ADF and PP tests. However, after applying the first difference, both ADF and PP tests reject the null hypothesis. As the data is found to be stationary after performing ADF and PP tests in first differences, no further tests was performed. In this study we apply stationary tests with trend effects. Therefore, the null hypothesis is rejected in both ADF and PP tests at $I(0)$:

Table 1 Stationary Test with Trend for ASEAN Tourist Arrival

<table>
<thead>
<tr>
<th></th>
<th>Level: $I(0)$</th>
<th>First Difference: $I(1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF test ($\tau$)</td>
<td>-1.096 (0)</td>
<td>-9.686 (0)*</td>
</tr>
<tr>
<td>PP test ($Z_{13}$)</td>
<td>-0.732 [13]</td>
<td>-11.094 [13]*</td>
</tr>
</tbody>
</table>

Note: Lag length in () and Newey-West value using Bartlett kernel in [ ]
Asterisks (*) denote statistically significant at 1% significance levels

Once stationarity is established, examination of the autocorrelation function plot (ACF) and partial autocorrelation function (PACF) plot over several quarterly lags suggests that autoregressive (AR) and moving average (MA) terms should be included in the ARIMA model. Therefore, we used ACF and PACF for tourist arrivals to Malaysia and choose the combination of ARIMA(p,d,q) to obtain the most suitable SARIMA model for this study. The standard procedure for identification, estimation, diagnostic checking and over fitting in a Box-Jenkins analysis of time series was performed. The estimation method involved maximum likelihood parameter estimation to obtain initial estimates and then unconditional least-squares estimation to obtain final estimates. It is a fairly common occurrence that differencing a time series introduces moving average terms into the resulting ARIMA model. Two criteria commonly applied to select between time series models are the AIC and SBC (Schwarz Bayesian Criteria). Both criteria evaluate the fit versus the number of parameters and in this study we used AIC as the main criteria to choose the best combination of ARIMA model. Table 2 indicates diagnostic correlogram with ACF and PACF for ARMA(2,2).
Table 2 Regression Results and Diagnostic Tests for SARIMA Models

<table>
<thead>
<tr>
<th>Models</th>
<th>Coefficient</th>
<th>ARCH-LM Test&lt;sup&gt;a&lt;/sup&gt;</th>
<th>H&lt;sub&gt;0&lt;/sub&gt;: No serial correlation&lt;sup&gt;b&lt;/sup&gt;</th>
<th>H&lt;sub&gt;0&lt;/sub&gt;: Normality Test&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>SARIMA(1,1,1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Season-I</td>
<td>0.08 (0.24)</td>
<td>3.91 (0.05)</td>
<td>0.93 (0.40)</td>
<td>23.63 (0.00)&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Season-II</td>
<td>-0.03 (0.64)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Season-III</td>
<td>0.08 (0.21)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.74 (0.00)&lt;sup&gt;*&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.99 (0.00)&lt;sup&gt;*&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC value</td>
<td>-0.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SARIMA(1,1,2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Season-I</td>
<td>0.09 (0.21)</td>
<td>2.21 (0.14)</td>
<td>1.12 (0.33)</td>
<td>31.99 (0.00)&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Season-II</td>
<td>-0.02 (0.68)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Season-III</td>
<td>0.09 (0.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.81 (0.00)&lt;sup&gt;*&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA(1)</td>
<td>-1.13 (0.00)&lt;sup&gt;*&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA(2)</td>
<td>0.14 (0.45)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC value</td>
<td>-0.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARIMA(2,1,2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Season-I</td>
<td>0.08 (0.39)</td>
<td>1.11 (0.29)</td>
<td>1.48 (0.24)</td>
<td>42.14 (0.00)&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Season-II</td>
<td>-0.02 (0.68)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Season-III</td>
<td>0.08 (0.40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(1)</td>
<td>-0.14 (0.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AR(2)</td>
<td>0.56 (0.00)&lt;sup&gt;*&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA(1)</td>
<td>-0.24 (0.38)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MA(2)</td>
<td>-0.74 (0.01)&lt;sup&gt;*&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC value</td>
<td>-0.68</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: (a), (b) and (c) indicates Autoregressive conditional heteroscedasticity (ARCH) LM test, Breusch-Godfrey (BG) serial correlation test and Jarque-Bera (JB) normality test. Figures in ( ) indicates probability values. Asterisks (*) denote statistically significant at 1% significance levels. Season-1 (April-June), Season-2 (July-September), Season-3 (October-December)

The flow of ACF and PACF shows the autoregressive effects with first difference of the historical data. Table 2 indicates AIC values and the decision to select the most suitable model is by comparing the value of AIC to the ARIMA models used in this study. The smaller the value of AIC, the better is the fit in the ARIMA model used. Therefore ARIMA(2,1,2) is relevant because the value of AIC is smaller compared to ARIMA models. Besides that, diagnostics tests are applied in this study to determine whether the estimated models deviate from the assumptions.
of standard linear regression model. Further, we tested for autoregressive conditional heteroscedasticity (ARCH), serial correlation using Breusch-Godfrey (BG) test; and normality using Jarque-Bera (JB) test. As correlogram of squared residuals from ARMA(2,2) shows autocorrelation pattern in squared residuals which could be attributed to volatility clustering, to test the presence of ARCH effect, we compute ARCH-LM test.

The results in Table 2 do not indicate any ARCH effects in all models estimated. The Breusch-Godfrey (BG) test of serial correlation indicates that serial correlation hypothesis cannot be rejected in all three SARIMA models. Meanwhile, normality test using Jarque-Bera (JB) indicates that normality in the errors is rejected in all SARIMA models. This indicates that all three models are not normally distributed because of some seasonal effects. One interesting results that can be derived from Table 2 is that, although seasonal effects have been taken into account in every SARIMA models, the results does not display any significant seasonal dummy effects, either positively or negatively.

To choose suitable SARIMA model for this study, we use the inequality coefficients technique. The inequality coefficient of SARIMA(2,1,2) model are marginally smaller than those of the SARIMA(1,1,1) and SARIMA(1,1,2). Therefore ARIMA(2,1,2) is the best SARIMA model for this study. For one-step-ahead forecasting which is applied in this study, the MA(2) model is found to be the best because the values of RMSE and MAPE are the lowest as in Table 3 below.

<table>
<thead>
<tr>
<th>Table 3 Summary of Forecast Evolutions of SARIMA Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inequality Coefficient</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td>RMSE</td>
</tr>
<tr>
<td>MAE</td>
</tr>
<tr>
<td>MAPE</td>
</tr>
<tr>
<td>Theil Coefficient</td>
</tr>
</tbody>
</table>

It is clear that the dummy variables which represent seasonality effects do not give any clear evidence for ASEAN tourist arrivals to Malaysia. Therefore, in order to forecast ASEAN tourist arrivals using one-period-ahead approach, this study used only ARIMA(2,1,2) model without any seasonal effect to identify the autoregressive (AR) and moving average (MA) effects. The ARIMA(2,1,2) model that had been
estimated for the sample period can be described as shown in the following equation with standard errors and t-values of the coefficients given in parentheses. Estimated AR(2) and MA(2) are found to be significant at 1%. It is clear that the AR(2) and MA(2) components are significant without any seasonal dummies because the dummy variables do not determine ASEAN tourist arrivals to Malaysia:

\[
\Delta \ln Tour_t = 0.02 - 0.22\Delta \ln Tour_{t-1} + 0.68\Delta \ln Tourist_{t-2} - 0.08\varepsilon_{t-1} - 0.91\varepsilon_{t-2}
\]

\[
t-stat \quad (7.74)^* \quad (-1.79) \quad (5.93)^* \quad (-0.84) \quad (-9.77)^*
\]

It worth mentioning that once we accept ARIMA(2,1,2) as a suitable model for this study, the model is used for forecasting purpose. We applied ARIMA(2,1,1) to forecast one-period ahead using historical data 1995:Q1 to 2009:Q4 and forecast the short-term period 2010:Q1 to 2010:Q4 of tourist arrivals to Malaysia. Figure 4 shows the forecasted ASEAN tourist arrivals to Malaysia using ARIMA(2,1,2) with one-period ahead forecasting method. Usage of quarterly data with short term forecasting, one-period-ahead procedure provides a slightly better forecast for Malaysia.

![Figure 5 One-Period Ahead Forecasted ASEAN Tourist Arrival (2010:Q1- 2010:Q4)](image)

**CONCLUSION**

378
To conclude, this study found that SARIMA(2,1,2) is not able to capture seasonality effects in predicting ASEAN tourist arrivals to Malaysia because seasonality does not have an effect on the numbers of ASEAN tourist arrivals to Malaysia and there is only autoregressive and moving average effects that appeared using ARIMA(2,1,2). The empirical forecasting method used in this study produces best fit ARIMA and SARIMA model and from a planning perspective this should be a major research theme in the study of international tourism demand. Further, incorporation of forecasts into decision making processes would assist development and investment strategies in tourism industry in the future.

REFERENCES


**SUBMITTED: JAN 2011**
**REVISION SUBMITTED: MAY 2011**
**ACCEPTED: JUN 2011**
**REFEREED ANONYMously**

**Nanthakumar Loganathan** (kumar@umt.edu.my) is a Senior Lecturer at Department of Economics, Faculty of Management and Economics, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia.
Thirunaukarasu Subramaniam (stkarasu@um.edu.my) is a Senior Lecturer at Department of Southeast Asian Studies, Faculty of Arts and Social Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia.

Mori Kogid (edy@ums.edu.my) is a Lecturer at School of Business and Economics, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia.