Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
A fractionally integrated autoregressive moving average approach to forecasting tourism demand

Fong-Lin Chu*

Graduate Institute of National Development, College of Social Science, National Taiwan University, No. 1, Roosevelt Road, Sec. 4, Taipei, Taiwan

Received 11 June 2005; accepted 2 February 2006

Abstract

The primary aim of this paper is to incorporate fractionally integrated ARMA \((p, d, q)\) (ARFIMA) models into tourism forecasting, and to compare the accuracy of forecasts with those obtained by previous studies. The models are estimated using the volume of monthly international tourist arrivals in Singapore. Empirical findings demonstrate the evidence that the approach we propose generates relatively lower sample mean absolute percentage errors (MAPEs). This study also deals with the volatile data faced by a forecaster. We use the Asian financial crisis and the September 11 event as examples of economic and political shocks. With respect to the objective of shaving the coefficient MAPE, forecasts based on the selected ARFIMA models dominate convincingly.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Fractionally integrated ARMA \((p, d, q)\); Forecast; Singapore; Tourist arrivals

1. Introduction

International tourism has experienced an overwhelming boom over the last two decades and has now been called the largest industry in the world. Despite this positive outcome the rate of growth has varied unevenly from year to year. While most countries have enjoyed a steady increase in their tourist arrivals, some have experienced a decline in numbers at times. Planning under these circumstances is exceptionally difficult and important. To be successful, accurate forecasts of tourism demand are required. This can avoid the financial costs of excess capacity or the opportunity costs of unfulfilled demand.

The end of World War II, as well as the development of relatively fast, safe and comfortable air travel, marked the genesis of modern-day international tourism. Slightly more than two-quarters of a century ago, the serious study of international tourism had begun. By virtue of the accumulating record of international travel statistics at this time, researchers began to study the international tourism demand, one area of which has been forecasts.

Over the last three to four decades, a substantial number of such studies have been undertaken. A large portion of these have yielded empirical information and a considerable majority have reported the accuracy of forecasts. Forecasting methods are generally classified into quantitative and qualitative (Archer, 1980; Uysal & Crompton, 1985). The quantitative approach, which gives more accurate forecasts than judgement forecasts are further divided into causal methods (e.g., regression and structural models) and time-series methods. The latter identify stochastic components (i.e., autoregressive and moving average components) in each time series and its distinguishing feature is that no attempt is made to relate a variable, say \(y\), to other variables. The movements in \(y\) are 'explained' solely in terms of its own past, or by its position in relation to time (Harvey, 1981).

The literature on modelling and forecasting tourism is huge. It includes approaches based on regression models, which to a large extent had been examined by Leob (1982), Uysal and Crompton (1985), Witt and Martin (1987), Crouch, Schultz, and Valerio (1992), Admowicz (1994), Morris, Wilson, and Bakalis (1995), Malenberg and Van Soest (1996), Kulendran and King (1997), and Song and Witt (2006). Other authors such as Geurts and Ibrahim

* Tel.: +886 2 33663326; fax: +886 2 23676176.
E-mail address: flchu@ntu.edu.tw.
(1975), Choy (1984), Van Doorn (1984), Martin and Witt (1989), Chan (1993), Chan, Hui, and Yuen (1999), Witt, Witt, and Wilson (1994), Turner, Kulendran, and Pergat (1995), Turner, Kulendran, and Fernando (1997), Garcia-Ferrer and Queralt (1997), Chu (1998a, b), Kim (1999), Lim and McAleer (2001, 2002), Goh and Law (2002), Gustavsson and Nordström (2001), Brännäs, Hellström, and Nordström (2002), Chu (2004), and Gil-Alana (2005) use pure time-series analytical models. In testing the accuracy of different forecasting models for tourist arrivals, researchers had found that time-series models often generate acceptable forecasts at low cost with reasonable benefits. Therefore, a spate of literature had investigated the forecasting ability along this line. Since the purpose of this paper is mainly concerned with the worldwide visitors in Singapore, a brief review of the articles dealing with the prediction of this island state’s tourism demand will be conducted. Chan (1993) examined various forecasting models on monthly data between July 1977 and December 1988 and found that the sine wave nonlinear regression model outperformed Naïve I, Naïve II, simple linear regression, and ARIMA (2, 1, 2). The mean absolute percentage error (MAPE) of the 19-months-ahead forecast is 2.567. Applying a seasonal ARIMA model and using the same dataset and forecasting horizon, Chu (1998a) found that the MAPE reduced to 1.857. Chu (1998b) further improved the MAPE to 1.813 by using a combined approach to the same dataset. This paper tries to improve the accuracy of forecasts by employing a fractionally integrated autoregressive moving average (ARFIMA) model. Sutcliffe (1994) argued that the use of fractional differencing opens up a much wider and realistic behavior for the trend and seasonal components than traditional integer differencing. He showed several advantages of using fractional differencing for forecasting monthly data. Franses and Ooms (1997) extended the standard ARFIMA (0, d, 0) model for quarterly UK inflation, where they allowed the fractional integration parameter d to vary with the season s. For quarterly out-of-sample forecasting, this so-called periodic ARFIMA (0, d, 0) model does not appear to be the overall winner when compared with rival forecasting schemes. Porter-Hudak (1990) characterized a set of monetary aggregate series using a seasonal ARFIMA model. She estimated the seasonal fractional differencing parameter using an extension of a method proposed by Geweke and Porter-Hudak (1983) for the non-seasonal fractional differencing case. She also looked at forecasts generated by the seasonal fractional model and compared the one-step-ahead forecasts with those of a standard seasonal ARIMA model. Ray (1993) used a series of monthly IBM product revenues to illustrate the usefulness of seasonal fractionally differenced ARMA models for business forecasting. She found that the fractionally differentiated seasonal model gives a more accurate next-quarter, next-half-year, next-year, and next-two-year forecasts than the non-fractional model based on the criteria that are specifically constructed to reflect the accuracy of long-range periodic forecasts. Applying a seasonal fractional integration to the number of monthly visitors to the US for the time period January 1996–February 2003, Gil-Alana (2005) shows that the series can be well described in terms of a seasonal I(d) process, with d higher than 0 but smaller than 1. To examine the forecasting properties in the same article, Gil-Alana used a longer dataset (the short-term arrivals at the Auckland International Airport, April 1978–December 1999) and found that the seasonal fractional models outperform the nonseasonal ones in practically all cases. However, existing literature regarding forecasting tourism demand for Singapore so far had not adopted ARFIMA processes and this paper will fill this gap. Monthly data from July 1977 to December 1988 are used to identify a model for the series. Immediately after that, a 19-months-ahead forecast is implemented. This is a time period that covers relatively stable data. The reason why a period of 19 months was selected as the forecasting horizon is to compare our results with that of previous studies. Subsequently, the dataset will be extended to account for economic and political shocks. Shocks normally generate relatively unstable data. Therefore, the model will be tested by relatively stable and unstable data. Using the Gulf War as an example of sudden environment change, Chan et al. (1999) had explored the relative performance of various forecasting models and found Naïve II to be the best in terms of forecasting accuracy. Chan, however, did not consider the ARFIMA model and restricted their prediction interval to 12 months. This paper considers ARFIMA and extends the sample data to cover the Asian financial crisis and the September 11 event. The prediction interval is extended to 19 months as well. The organization of this paper is as follows. The next section briefly discusses Singapore’s tourist industry and its time-series data. The third section presents the ARFIMA model. The fourth section analyzes model estimation and reports the application to tourist arrivals forecasting. The final section presents a summary and conclusions.

2. Singapore tourism industry and data

The performance of the Singapore tourism industry has been remarkable in recent years. The island state’s tourism sector, one of its pillar industries, employs some 150,000 people and accounts for about 10% of its economy of $87 billion a year. It also attracts some 8 million visitors—twice its population—each year from all over the world. The data of monthly number of international tourists visiting Singapore between July 1977 and November 2004 are obtained from Singapore Tourist Promotion Board and are plotted in Fig. 1. The marked seasonal pattern dominates the movement in the series. The plot exhibited a permanent deterministic pattern of a long-term upward trend with short-term fluctuations that were independent from one time period to the next. Note that Singapore’s popularity has been growing; the series appears to be nonstationary in
that the mean is increasing over time. The tourism industry survived the heavy blows of the first Iraq War, the Asian financial crisis, the September 11 terrorist attacks on the United States in 2001, the subsequent Iraq War, and the SARS (severe acute respiratory syndrome) outbreak in early 2003. These events, however, had influenced the decision of many of the potential travelers, and tourist arrivals therefore had been reduced significantly below the linear trend line in Fig. 1. Among them, the Asian financial crisis and the September 11 event are two different demand shocks in nature. The former is economic while the latter is political. The impacts of both disturbances on arrivals are demonstrated by a fenced rectangular in Fig. 1. For the forecasting experiment, the sample was split into the last 19 observations, henceforth called the ‘prediction interval’, and the remaining data, henceforth called the ‘pre-prediction interval’. Parameter estimates were obtained from fitting the models to the pre-prediction interval. Three different types of pre-prediction and prediction interval are considered. One of them has stable tourist arrivals within the entire sample, henceforth called the ‘tranquil period’. The other two, characterized by economic and political shocks, are further classified into cases, depending upon whether shock has taken place at the time the policy maker or the planning manager implemented the forecasts. Suppose the policy maker had already implemented the forecasts based on his best forecasting model but found later that shock had occurred. In this situation, the policy maker may be reluctant to change the forecasts because cost may surpass benefit. Even if he does change the forecasts based on alternative models at hand, there is no guarantee that the results will be better than those obtained from the previous models. Alternatively, if the shock had already taken place at the time the forecast is made, but the industry still shows no sign of partial or full recovery, the policy maker in this situation has the option to try alternative forecasting models if he deems appropriate.

The Asian financial crisis took place in the second half of 1997. Accordingly, prediction intervals consist of January 1997–July 1998 and January 1998–July 1999. In the former, demand shock was not present at the time the forecast was made. In the latter, shock had already taken place. Parameter estimates were obtained from fitting the models to the pre-prediction interval, which corresponds to July 1977–December 1996 and July 1977–December 1997, respectively. The September 11 event took place in the latter half of 2001. Likewise, the prediction intervals are January 2001–July 2002 and January 2002–July 2003. Pre-prediction and prediction intervals for various cases are summarized in Table 1.

### Table 1

| Pre-prediction and prediction interval for various stages | 
|--------------------------------------------------------|
| **Pre-prediction interval** | **Prediction interval** |
| **Tranquil period** | 
| July 1977–December 1988 | January 1989–July 1990 |
| **Asian financial crisis** | 
| July 1977–December 1996 | January 1997–July 1998 |
| July 1977–December 1997 | January 1998–July 1999 |
| **September 11 event** | 
| July 1977–December 2000 | January 2001–July 2002 |
| July 1977–December 2001 | January 2002–July 2003 |

### 3. A fractionally integrated ARMA model

A long memory model, which has become popular in econometrics, is the (discrete time) fractionally differenced model (Brockwell & Davis, 2002). It was suggested as an alternative to ARIMA models by Hosking (1981), and Granger and Joyeux (1980). A fractionally integrated ARMA processes, or more precisely, the ARIMA \((p, d, q)\) processes with \(0 < d < 0.5\), satisfy difference equations of the form

\[
(1 - B)^d \phi(B)X_t = \theta(B)Z_t, \tag{1}
\]

where \(\phi(z)\) and \(\theta(z)\) are polynomials of degrees \(p\) and \(q\), respectively, satisfying \(\phi(z)\neq0\) and \(\theta(z)\neq0\) for all \(z\) such that \(|z|\leq1\), \(B\) is the backward shift operator (i.e., \(BX_t = X_{t-1}\)) and \(\{Z_t\}\) is a Gaussian white noise sequence with mean 0 and variance \(\sigma^2\). The difference operator \((1-B)^d\) is raised to a fractional power, \(d\), denoting the **fractional order of integration**. The operator \((1-B)^d\) is defined by the binomial expansion

\[
(1 - B)^d = \sum_{j=0}^{\infty} \pi_j B^j,
\]

where

\[
\pi_j = \prod_{0 < k < j} \frac{k - 1 - d}{k}, \quad j = 1, 2, \ldots
\]
The autocorrelations \( \rho(h) \) at lag \( h \) of an ARIMA(\( p, d, q \)) process with \( 0 < d < 0.5 \) has the property
\[
\rho(h)h^{-2d} \rightarrow c \quad \text{as} \quad h \rightarrow \infty. \tag{2}
\]
This hyperbolic decay at high lags distinguishes series with long memory from series with short memory. It is the main characteristic for empirical identification.

The spectral density of \( \{X_t\} \) is given by
\[
f(\lambda) = \frac{\sigma^2}{2\pi} \left| \theta(e^{-i\lambda}) \right|^2 \left| 1 - e^{-i\lambda} \right|^{-2d}. \tag{3}
\]

Calculation of exact Gaussian likelihood of observations \([x_1, \ldots, x_n]\) of a fractionally integrated ARMA process is very slow and demanding in terms of computer memory. Instead of estimating the parameters \( d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q \), and \( \sigma^2 \) by maximizing the exact Gaussian likelihood, it is simpler to maximize the Whittle approximation \( L_w \), defined by
\[
-2 \ln (L_w) = n \ln (2\pi) + 2n \ln \sigma
+ \sigma^{-2} \sum_j I_n(\omega_j) + \sum_j \ln g(\omega_j), \tag{4}
\]
where \( I_n \) is the periodogram \( \sigma^2 g/(2\pi) \) is the model spectral density, and \( \sum_j \) denotes the sum over all nonzero Fourier frequencies \( \omega_j = 2\pi/n \in (-\pi, \pi) \). The parameters of the fractionally integrated ARIMA(\( p, d, q \)) process can be estimated and the results can be used for out-of-sample forecasting evaluation.

### 4. Model estimation and forecasting

Because the series illustrated in Fig. 1 demonstrates a combination of strong seasonality and nonstationarity, they are 12th differenced in an effort to obtain the seasonally differenced data. In order to impart stationarity into the series, we next take the first difference of the already seasonally differenced data. The results are displayed in Fig. 2. Alternatively, one may just take the first difference of the raw data and ignore the 12th difference. This is illustrated in Fig. 3. Sometimes forecasts based on the model fit to first differenced data render better results.

![Fig. 3. First differenced data.](image)

### 4.1. Empirical results for the tranquil period

The following models, each with different parameters, are found to provide a good fit to the data for the tranquil period. The first model considered was the parsimonious \((0, d, 1)\) structure as suggested by Akaike information criterion (AIC):
\[
(1 - B)^{0.098}(\Delta^d X_t) = \theta(B)Z_t,
\theta(B) = 1 - 0.592B. \tag{5}
\]
The \( dth \) difference of a series is denoted by \( \Delta^d \). A seasonal difference is denoted by \( \Delta_s \), where \( s \) is the period of the data. The \( D \)th such seasonal difference is \( \Delta_s^D \). In Eq. (5), we use monthly data and therefore 12th difference the data first and then first difference the resulting series, which is denoted by \( \Delta^1 \Delta_s^{12} X_t \). The second best model suggested by AIC was the \((4, d, 6)\) structure:
\[
\phi(B)(1 - B)^{0.269}(\Delta^4 X_t) = \theta(B)Z_t,
\phi(B) = 1 + 0.716B + 0.144B^2 + 0.297B^3 + 0.618B^4,
\theta(B) = 1 - 0.041B - 0.643B^2 + 0.103B^3 + 0.851B^4
- 0.214B^5 - 0.550B^6. \tag{6}
\]

Note that the resulting structure in Eq. (5) took into account the seasonal difference, while Eq. (6) ignored it. The forecasting performance of both models will be compared later. A frequently used descriptive device for measuring forecasting performance is the sample MAPE (Makridakis & Hibon, 1979; Makridakis et al., 1984). The coefficient MAPE measures the mean of absolute percent error of prediction and is given by
\[
\text{MAPE} = \frac{1}{K} \sum_{i=1}^{N+K} \left| 1 - \frac{\hat{X}_t}{X_i} \right| \cdot 100,
\]
where \( \hat{X}_t \) and \( X_i \) represent the forecast and actual values. The MAPEs for Eqs. (5) and (6) were 1.738 and 7.578%, respectively. Guided by AIC and out-of-sample analysis of forecasting accuracy, we choose the structure of Eq. (5) for the tranquil period. The results are illustrated in Table 2 along with the outcomes of previous empirical studies in an attempt to make a comparison.

Table 2 illustrates the forecasting ability of nine time-series models, in which the fractionally integrated ARMA approach turned out to be the best. It is notoriously
difficult to produce a value of MAPE less than 2%, as far as the ex post forecasting of monthly tourist arrivals is concerned. This is especially so when the prediction horizon lasts for 19 months. Among the models which rendered a MAPE of less than 2%, the fractionally integrated ARMA approach has a 4.15% and 6.41% improvement over the combined forecast and ARIMA (3, 1, 0)(0, 1, 0)\textsubscript{12}, respectively. The improvement over the rest of the models is even more remarkable. All pose double-digit percentage MAPE improvement. The sample forecasts and actual arrivals are graphically illustrated in Fig. 4. The forecasts follow the overall trend in tourist arrivals, and do capture the cyclical fluctuations that occurred during the prediction interval. The forecasts reproduce most of the turning points as well.

Table 3 shows both the forecasts and actual numbers. For comparison among the models which manufactured MAPEs below 2%, the results of ARIMA (3, 1, 0) (0, 1, 0)\textsubscript{12} and the combined forecast models are also presented. Forecasts with minimal errors are in bold face. The mean and standard deviation of the absolute percentage errors are illustrated in Table 4. Regarding the overall performances, the selected ARFIMA dominated convincingly.

4.2. The period of the Asian financial crisis and the September 11 event

Next we proceed to analyze the choice of the models pertaining to the period of the Asian financial crisis. At first we look at situation where economic shock only hit the prediction interval. In other words, when the policy maker implements the prediction task, shock has not yet taken place. Subsequently, shock did indeed hit the prediction interval so that the policy maker was dumbfounded and could do nothing but wait for the actual arrivals. The first model considered was the parsimonious (2, d, 2) structure as suggested by AIC:

\[
\phi(B)(1 - B)^{-0.411}(\Delta^4 \Delta_{12}X_t) = \theta(B)Z_t, \\
\phi(B) = 1 + 1.026B + 0.832B^2, \\
\theta(B) = 1 + 1.042B + 0.982B^2.
\]  

Fig. 4. Actual and forecasted number of tourist arrivals.

Next we proceed to analyze the choice of the models pertaining to the period of the Asian financial crisis. At first we look at situation where economic shock only hit the prediction interval. In other words, when the policy maker implements the prediction task, shock has not yet taken place. Subsequently, shock did indeed hit the prediction interval so that the policy maker was dumbfounded and could do nothing but wait for the actual arrivals. The first model considered was the parsimonious (2, d, 2) structure as suggested by AIC:

\[
\phi(B)(1 - B)^{-0.411}(\Delta^4 \Delta_{12}X_t) = \theta(B)Z_t, \\
\phi(B) = 1 + 1.026B + 0.832B^2, \\
\theta(B) = 1 + 1.042B + 0.982B^2.
\]  

The second best model suggested by AIC was the (4, d, 6) structure:

\[
\phi(B)(1 - B)^{0.126}(\Delta^4 \Delta X_t) = \theta(B)Z_t, \\
\phi(B) = 1 + 1.744B + 1.988B^2 + 1.654B^3 + 0.936B^4 - 0.865B^5 - 0.227B^6, \\
\theta(B) = 1 + 1.361B + 0.907B^2 - 0.087B^3 - 0.732B^4 - 0.481B^5 - 0.087B^6.
\]  

The calculated sample MAPEs for Eqs. (7) and (8) are 14.40% and 11.86%, respectively. Guided by the out-of-sample analysis of forecasting accuracy, we choose the structure of Eq. (8). For the situation where the policy
makers face a pre-prediction interval in which economic shock has already taken place, the first model considered was the parsimonious \((1, d, 1)\) structure as suggested by AIC:

\[
\phi(B)(1 - B)^{-0.213}(\Delta^1 X_t) = \theta(B)Z_t,
\]
\[
\phi(B) = 1 + 0.665B,
\]
\[
\theta(B) = 1 + 0.432B.
\]

The second best model suggested by AIC was the \((11, d, 12)\) structure:

\[
\phi(B)(1 - B)^{-0.061}(\Delta^1 X_t) = \theta(B)Z_t,
\]
\[
\phi(B) = 1 + 0.619B + 0.715B^2 + 0.454B^3 + 0.184B^{11},
\]
\[
\theta(B) = 1 + 0.303B + 0.438B^2 + 0.447B^{12}.
\]

The calculated sample MAPEs for Eqs. (9) and (10) are 8.79\% and 6.36\%, respectively. Guided by the out-of-sample analysis of forecasting accuracy, we choose the structure of Eq. (10).

Preferred structures pertaining to the period of the September 11 event are the following alternative models. This is the situation where political shock only hit the prediction interval. The first model considered was the parsimonious \((1, d, 0)\) structure as suggested by (AIC):

\[
\phi(B)(1 - B)^{-0.179}(\Delta^1 X_t) = Z_t,
\]
\[
\phi(B) = 1 + 0.252B.
\]

The second best model suggested by AIC was the \((2, d, 5)\) structure:

\[
\phi(B)(1 - B)^{-0.128}(\Delta^1 X_t) = \theta(B)Z_t,
\]
\[
\phi(B) = 1 + 0.008B + 0.999B^2,
\]
\[
\theta(B) = 1 - 0.319B + 1.057B^2 - 0.700B^3
\]
\[
+ 0.174B^4 - 0.323B^5.
\]

The MAPEs for Eqs. (11) and (12) are 14.69\% and 4.71\%, respectively. Guided by the out-of-sample analysis of forecasting accuracy, we choose the structure of Eq. (12). In the situation where political shock hit the pre-prediction interval, the first model considered was the parsimonious \((1, d, 0)\) structure as suggested by AIC:

\[
\phi(B)(1 - B)^{-0.202}(\Delta^1 X_t) = Z_t,
\]
\[
\phi(B) = 1 + 0.229B.
\]

The second best model suggested by AIC was the \((3, d, 6)\) structure:

\[
\phi(B)(1 - B)^{-0.277}(\Delta^1 X_t) = \theta(B)Z_t,
\]
\[
\phi(B) = 1 + 0.622B + 1.002B^2 + 0.617B^3,
\]
\[
\theta(B) = 1 + 0.562B + 0.955B^2 + 0.301B^3 - 0.360B^4
\]
\[
- 0.133B^5 - 0.373B^6.
\]

The MAPEs for Eqs. (13) and (14) are 30.82\% and 31.15\%, respectively. Guided by AIC and the out-of-sample analysis of forecasting accuracy, we choose the structure of Eq. (13).

The ex post forecasts for the period of the Asian financial crisis and the September 11 event are generated. MAPEs are calculated and illustrated in Tables 5 and 6, respectively. For comparison, parameter estimates are obtained from fitting the various models to the pre-prediction interval. These include univariate extrapolative models such as simple linear regression, Naive I, Naive II, cubic polynomial, and sine wave nonlinear regression. Modern stochastic time-series model such as nonseasonal ARIMA \((p, d, q)\) and seasonal ARIMA models are also considered. Model-based out-of-sample forecasts are obtained and the MAPEs are calculated and displayed in Tables 5 and 6. The minimal MAPE manufactured by a specific model is illustrated by bold-faced numerical values.

Again the results in Tables 5 and 6 seem to indicate that the fractionally integrated ARMA approach relatively outperformed the rest of the models. It generated the best predictor for both prediction intervals during the Asian financial crisis. It also generated the best and the second best predictor for both prediction intervals during the September 11 event. The result for the prediction interval of January 2001–July 2002 is worth mentioning. Here, the coefficient MAPE is merely 4.71\%, despite the tourist flows being influenced by the September 11 event. The deterioration of MAPEs seen in the prediction interval of January 2002–July 2003 mostly reflected the plunge of international

| Table 4 |
| --- |
| Mean and standard deviation of absolute percentage errors of various models |
| Model | Mean | Std. deviation |
| --- | --- | --- |
| Seasonal ARIMA | 1.857\% | 1.260 |
| Combined forecast | 1.813 | 1.124 |
| Fractionally integrated ARMA | 1.738 | 0.951 |

| Table 5 |
| --- |
| Sample MAPE of various models for the period of the Asian financial crisis |
| Model | Prediction interval |
| --- | --- |
| Naive I | 1/97–7/98 1/98–7/99 |
| Naive II | 20.64\% 8.89\% |
| Linear regression | 36.20 16.34 |
| Cubic polynomial regression | 12.95 19.62 |
| Sine wave nonlinear regression | 14.03 7.31 |
| ARIMA \((p, d, q)\) | 13.05 20.97 |
| ARIMA \((p, d, q)(P, D, Q)_{12}\) | 14.30 6.55 |
| Fractionally integrated ARIMA \((p, d, q)\) | 11.86 6.36 |

\(^{a}ARIMA \((2, 1, 2)\).\)
\(^{b}ARIMA \((2, 1, 2)\).\)
\(^{c}ARIMA \((1, 1, 1)(0, 1, 1)_{12}\).\)
\(^{d}ARIMA \((1, 1, 2)(0, 1, 1)_{12}\).\)
In all cases, fractionally integrated ARIMA did better than the cubic polynomial regression method. The prediction interval had not been severely disturbed by the integrated ARMA. In the best outcome, the cubic polynomial regression seems to produce more visitors when the SARS crisis was at its peak. For that period of tranquility, the Asian financial crisis, and the September 11 event, the cubic polynomial regression did best. For SARS, the SARS crisis was at its peak. For that period of 19 months, the SARS crisis was at its peak. For that period of tranquility, the Asian financial crisis, and the September 11 event. The minimal MAPE manufactured by a specific model is illustrated by bold-faced numerical values. Indeed, the fractionally integrated ARMA outperformed the rest of the models in most of the cases.

To this end, forecasting comparison was done statically, i.e., the models were fitted with the data up to a certain time, and then the estimated models were used to forecast the tourist arrivals over the next 12, 19 and 24 months. An alternative way for the forecasting assessment is to calculate the MAPEs recursively and then compare the forecasting performance with different forecasting horizons. Here, 1-2- and 6-year out-of-sample forecasts are implemented. The parameters of each model are re-estimated at the beginning of each forecast period using all of the observations up to the forecast origin. Tables 7-9 illustrate the MAPEs of various models for the period of tranquility, the Asian financial crisis, and the September 11 event, respectively. The minimal MAPE manufactured by a specific model is illustrated by bold-faced numerical values. Indeed, the fractionally integrated ARMA outperformed the rest of the models in most of the cases.

Forecasting performance of the tourism demand models tend to vary with the nature of the data used, i.e., monthly series, viz., quarterly series. To ensure the reliability of the models, we fit all models and make 12- and 24-month forecasts. Tables 10–12 give the MAPE statistics for the periods of tranquility, the Asian financial crisis, and the September 11 event, respectively. The forecast generated by the fractionally integrated ARMA model appears to predict the monthly arrivals quite well except those generated during the months when the SARS crisis took place.

Table 6 Sample MAPE of various models for the period of the September 11 event

| Model                  | Prediction interval | 1/01–7/02 | 1/02–7/03 |
|------------------------|---------------------|-----------|-----------|
| Naïve I                | 11.28%              | 39.16%    |           |
| Naïve II               | 23.03               | 74.09     |           |
| Linear regression      | 8.81                | 43.74     |           |
| Cubic polynomial regression | 7.78               | 29.86     |           |
| Sine wave nonlinear regression | 8.90               | 43.73     |           |
| ARIMA (p, d, q)        | 6.84a               | 33.73b    |           |
| ARIMA (p, d, q)(P, D, Q) | 8.46c               | 31.54d    |           |
| Fractionally integrated ARIMA (p, d, q) | 4.71               | 30.82     |           |

Notes: a: ARIMA (2, 1, 3). b: ARIMA (4, 1, 3). c: ARIMA (3, 1, 3)(1, 0, 1)12. d: ARIMA (0, 0, 1)(1, 1, 1)12.

Table 7 Sample MAPE of various models for 12- and 24-months-ahead forecasts

| Model                  | Prediction interval | 1/01–7/02 | 1/02–7/03 |
|------------------------|---------------------|-----------|-----------|
| Naïve I                | 7.42%               | 9.17%     |           |
| Naïve II               | 8.81                | 6.76      |           |
| Linear regression      | 16.16               | 18.38     |           |
| Cubic polynomial regression | 2.11               | 5.77      |           |
| Sine wave nonlinear regression | 2.58               | 2.82      |           |
| ARIMA (p, d, q)        | 7.83                | 11.12     |           |
| ARIMA (p, d, q)(P, D, Q) | 1.74               | 2.03      |           |
| Fractionally integrated ARIMA (p, d, q) | 1.63               | 2.28      |           |

Table 8 Sample MAPE of various models for 12- and 24-months-ahead forecasts

| Model                  | Prediction interval | 1/01–7/02 | 1/02–7/03 |
|------------------------|---------------------|-----------|-----------|
| Naïve I                | 14.12%              | 22.47%    | 11.17%    | 8.12%    |
| Naïve II               | 28.77               | 38.32     | 19.69     | 13.79    |
| Linear regression      | 7.81                | 15.52     | 22.65     | 18.52    |
| Cubic polynomial regression | 12.61              | 24.03     | 24.73     | 19.76    |
| Sine wave nonlinear regression | 8.15               | 15.72     | 2.96      | 19.85    |
| ARIMA (p, d, q)        | 9.14                | 14.76     | 8.63      | 7.29     |
| ARIMA (p, d, q)(P, D, Q) | 9.13               | 20.13     | 5.76      | 9.13     |
| Fractionally integrated ARIMA (p, d, q) | 7.19               | 13.18     | 5.29      | 7.48     |

Table 9 Sample MAPE of various models for 12- and 24-months-ahead forecasts

| Model                  | Prediction interval | 1/01–7/02 | 1/02–7/03 |
|------------------------|---------------------|-----------|-----------|
| Naïve I                | 11.79%              | 11.19%    | 7.65%     | 33.89%   |
| Naïve II               | 22.96               | 22.38     | 32.89     | 66.20    |
| Linear regression      | 8.25                | 9.59      | 9.63      | 38.75    |
| Cubic Polynomial regression | 6.99               | 8.12      | 5.85      | 25.15    |
| Sine wave nonlinear regression | 8.24               | 9.57      | 9.62      | 38.75    |
| ARIMA (p, d, q)        | 7.49                | 6.73      | 5.83      | 28.03    |
| ARIMA (p, d, q)(P, D, Q) | 7.23               | 9.03      | 3.73      | 26.25    |
| Fractionally integrated ARIMA (p, d, q) | 5.67               | 4.98      | 6.23      | 26.65    |
forecasting evaluation, our estimated models were further checked by quarterly time series in the forecasting exercise for all datasets. Guided by AIC and the out-of-sample analysis of forecasting accuracy, we choose the models represented by the following equations:

\[
\phi(B)(1 - B)^{0.360}(\Delta^1 X_t) = \theta(B)Z_t, \\
\phi(B) = 1 + 0.867B + 0.136B^2 + 0.091B^3 - 0.848B^4, \\
\theta(B) = 1 - 0.315B + 0.298B^2 - 0.039B^3 - 0.416B^4, \\
\]  

\[
\phi(B)(1 - B)^{-0.241}(\Delta^1 X_t) = \theta(B)Z_t, \\
\phi(B) = 1 - 0.014B + 0.057B^2 - 1.166B^4 + 0.239B^8 \\
- 0.015B^6, \\
\theta(B) = 1 - 0.249B + 0.41B^2 - 0.712B^4 - 0.097B^6 \\
+ 0.337B^8, \\
\]  

\[
\phi(B)(1 - B)^{-0.097}(\Delta^1 X_t) = \theta(B)Z_t, \\
\phi(B) = 1 + 0.032B + 0.9428B^4, \\
\theta(B) = 1 - 0.053B - 0.040B^2 - 0.546B^4, \\
\]  

\[
\phi(B) = 1 - 0.992B + 0.995B^2 + 0.994B^3, \\
\theta(B) = 1 + 0.657B + 0.767B^2 + 0.686B^3, \\
\]  

\[
\phi(B) = 1 + 0.001B - 0.008B^2 - 0.962B^4, \\
\theta(B) = 1 - 0.241B + 0.020B^2 - 0.719B^4. \\
\]  

The results of the calculated sample MAPEs are illustrated in Table 13. We produced forecasts for 4, 6, and 8 quarters and the results demonstrated that the forecasts generated by quarterly series perform as well as those of monthly series. For example, the forecasts generated for the tranquil period is so magnificent that the coefficient MAPE is approximately 2%.

5. Conclusions

A fractionally integrated ARMA (p, d, q) model has been constructed to explain the demand for Singapore tourism. The data for the inflow of international visitors over the periods July 1977–December 1988, July 1977–December 1996, July 1977–December 1997, July 1977–December 1998.
Table 12
Sample MAPE for forecasts of monthly tourism arrivals

| Forecast origin | Forecast period | MAPE  | Forecast origin | Forecast period | MAPE  |
|-----------------|----------------|-------|-----------------|-----------------|-------|
| 12/00           | 1/01           | 4.01  | 12/01           | 1/02            | 0.47  |
| 1/01            | 2/01           | 4.90  | 1/02            | 2/02            | 3.40  |
| 2/01            | 3/01           | 0.65  | 2/02            | 3/02            | 2.28  |
| 3/01            | 4/01           | 7.89  | 3/02            | 4/02            | 1.76  |
| 4/01            | 5/01           | 0.45  | 4/02            | 5/02            | 0.85  |
| 5/01            | 6/01           | 4.64  | 5/02            | 6/02            | 1.41  |
| 6/01            | 7/01           | 3.97  | 6/02            | 7/02            | 1.39  |
| 7/01            | 8/01           | 0.09  | 7/02            | 8/02            | 0.22  |
| 8/01            | 9/01           | 8.59  | 8/02            | 9/02            | 2.88  |
| 9/01            | 10/01          | 4.98  | 9/02            | 10/02           | 2.62  |
| 10/01           | 11/01          | 10.02 | 10/02           | 11/02           | 1.75  |
| 11/01           | 12/01          | 6.51  | 11/02           | 12/02           | 1.14  |
| 12/01           | 1/02           | 6.01  | 12/02           | 1/02            | 1.41  |
| 1/02            | 2/02           | 5.35  | 2/02            | 2/02            | 2.12  |
| 2/02            | 3/02           | 3.17  | 3/02            | 3/02            | 2.91  |
| 3/02            | 4/02           | 1.78  | 4/02            | 4/02            | 177.85|
| 4/02            | 5/02           | 0.95  | 5/02            | 5/02            | 37.93 |
| 5/02            | 6/02           | 17.56 | 6/02            | 6/02            | 2.83  |
| 6/02            | 7/02           | 1.61  | 7/02            | 7/02            | 2.33  |

2000, and July 1977–December 2001 are fitted to the mathematical model that we constructed. For each pre-

Table 13
Sample MAPE for forecasts of quarterly tourism demand

| Model          | Prediction interval | MAPE  |
|----------------|---------------------|-------|
| Eq. (15)       | 1Q/89–4Q/89         | 2.07% |
| Eq. (16)       | 1Q/89–2Q/90         | 2.16% |
| Eq. (17)       | 1Q/89–2Q/90         | 1.73% |
| Eq. (18)       | 1Q/97–4Q/97         | 6.43% |
| Eq. (19)       | 1Q/97–2Q/97         | 10.35 |
| Eq. (20)       | 1Q/97–2Q/97         | 12.85 |
| Eq. (21)       | 1Q/97–2Q/97         | 6.38% |
| Eq. (22)       | 1Q/97–2Q/97         | 5.02% |
| Eq. (23)       | 1Q/97–2Q/97         | 5.72% |
| Eq. (24)       | 1Q/97–2Q/97         | 6.58% |
| Eq. (25)       | 1Q/97–2Q/97         | 6.37% |
| Eq. (26)       | 1Q/97–2Q/97         | 5.83% |
| Eq. (27)       | 1Q/97–2Q/97         | 4.62% |
| Eq. (28)       | 1Q/97–2Q/97         | 28.56 |
| Eq. (29)       | 1Q/97–2Q/97         | 23.68 |

resulting parameter estimates to generate the ex post out-
of-sample forecasts of the number of arrivals for the next 19 months. MAPE was the criterion used to evaluate our forecasting performance. During the tranquil period, the fractionally integrated ARMA \( (p, d, q) \) model is the clear winner amongst the nine rival models. In fact, it improves the accuracy of forecast by several percentage points depending upon which rival models we are comparing it with. For instance, the improvement over the combined forecasting approach and seasonal ARMA \( (3, 1, 0)(0, 1, 0)_{12} \) model are 4.15% and 6.41%, respectively. Improvement over the rest of the models is even more magnificent.

The model we proposed also performed quite well in environments disturbed by economic and political shocks. Although we differentiate two cases that the forecaster or the policy maker may face in the vicinity of the shock, the proposed models seem to render the best performance to the policy maker among the methods we consider in most cases. This study is motivated by the potential of gains in predictive accuracy by using a fractionally integrated ARMA \( (p, d, q) \) model, in which the estimates of differencing parameter \( d \) are treated as fractional. Apart from treating \( d \) as a noninteger, appropriate integer differences are also implemented simultaneously in an effort to impart the data stationary. This may explain why the fractionally integrated ARMA \( (p, d, q) \) model performed relatively better than that of its rival traditional and seasonal ARIMA \( (p, d, q) \).

We also calculated the MAPEs recursively and then compared the forecasting performance with different forecasting horizons as an alternative approach for the forecasting assessment. One, two, and six periods ahead, forecasts are executed. The study was further expanded by including quarterly series in the forecasting exercise to ensure the reliability of the forecasting evaluation. Apart from the SARS period, the fractionally integrated ARMA \( (p, d, q) \) model outperforms the rival forecasting models most of the time.

The contribution of this paper is to show that a fractionally integrated ARMA model can be a credible alternative when modeling monthly tourism arrivals in Singapore. When model-based out-of-sample forecasts are entertained, more accurate results are generated. Hence, the results presented in this paper lead us to suggest that they can be used by recreation managers and decision makers in regional and national tourism offices in implementation of policies in an attempt to enhance tourism demand and growth.

An interesting topic for future research is to employ this estimating approach to international tourism demand for other countries. Model-based out-of-sample forecasts for the arrivals can also be implemented and used to find out whether the accuracy of forecast has improved. Tourism market in the West might possess characteristics different from those of the Asian Pacific countries. But when demand shocks take place, the global tourism industry can hardly escape being impacted. The need to thoroughly test...
the model inclusive of a static and disturbed environment is inevitable for a sound model.

References

Admowicz, W. L. (1994). Habit formation and variety seeding in a discrete choice model of recreation demand. *Journal of Agricultural and Resource Economics*, 19, 19–31.

Archer, B. H. (1980). Forecasting demand: Quantitative and intuitive technique. *International Journal of Tourism Management*, 1, 5–12.

Brännäs, K., Hellström, J., & Nordström, J. (2002). A new approach to modeling and forecasting monthly guest nights in hotels. *International Journal of Forecasting*, 18, 19–30.

Brockwell, P. J., & Davis, R. A. (2002). *Introduction to time series and forecasting* (2nd ed.). New York: Springer.

Chan, Y. M. (1993). Forecasting tourism: A sine wave time series regression approach. *Journal of Travel Research*, 32, 58–60.

Chan, Y. M., Hui, T. K., & Yuen, E. (1999). Modeling the impact of sudden environmental changes on visitor arrival forecasts: The case of the Gulf War. *Journal of Travel Research*, 37, 391–394.

Choy, D. L. (1984). Forecasting tourism revisited. *Tourism Management*, 5, 171–176.

Chu, F. L. (1998a). Forecasting tourism arrivals: Nonlinear sine wave or ARIMA? *Journal of Travel Research*, 36, 79–84.

Chu, F. L. (1998b). Forecasting tourism arrivals: A combined approach. *Tourism Management*, 19, 515–520.

Chu, F. L. (2004). Forecasting tourism demand: A cubic polynomial approach. *Tourism Management*, 25, 209–218.

Crouch, G. L., Schultz, L., & Valerio, P. (1992). Marketing international tourism to Australia: A regression analysis. *Tourism Management*, 13, 196–208.

Fransen, P. H., & Ooms, M. (1997). A periodic long-memory model for quarterly UK inflation. *International Journal of Forecasting*, 13, 117–126.

Garcia-Ferrer, A., & Queralt, R. A. (1997). A note on forecasting international tourism demand in Spain. *International Journal of Forecasting*, 13, 539–549.

Geurts, M., & Ibrahim, I. B. (1975). Comparing the Box–Jenkins approach with the exponentially smoothed forecasting model: Application to Hawaii tourists. *Journal of Marketing Research*, 12, 182–188.

Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4, 221–238.

Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, 23, 499–510.

Gil-Alana, L. A. (2005). Modeling international monthly arrivals using seasonal univariate long-memory processes. *Tourism Management*, 26, 867–878.

Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–39.

Gustavsson, P., & Nordström, J. (2001). The impact of seasonal unit roots and vector ARMA modeling on forecasting monthly tourism flows. *Tourism Economics*, 7, 117–133.

Harvey, A. C. (1981). *Time series models*. Deddington: Philip Allan.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68, 165–176.

Kim, J. H. (1999). Forecasting monthly tourist departure from Australia. *Tourism Economics*, 5, 277–291.

Kulendran, N., & King, M. L. (1997). Forecasting international quarterly tourist flows using error-correction and time-series models. *International Journal of Forecasting*, 13, 319–327.

Leob, P. D. (1982). International travel to the United States: An econometric evaluation. *Annals of Tourism Research*, 9, 7–20.

Lim, C., & McAleer, M. (2001). Cointegration analysis of quarterly tourism demand by Hong Kong and Singapore for Australia. *Applied Economics*, 33, 1599–1619.

Lim, C., & McAleer, M. (2002). Time series forecasts of international travel demand for Australia. *Tourism Management*, 23, 389–396.

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., & Lewandowske, R., et al. (Eds.). (1984). *The accuracy of major forecasting procedure*. New York: Wiley.

Makridakis, S., & Hibon, M. (1979). Accuracy of forecasting: An empirical investigation. *Journal of the Royal Statistical Society*, 142, 97–125.

Martin, C. A., & Witt, S. F. (1989). Forecasting tourism demand: A comparison of the accuracy of several quantitative methods. *International Journal of Forecasting*, 5, 7–19.

Malenberg, B., & Van Soest, A. (1996). Parametric and semiparametric modeling of vacation expenditures. *Journal of Applied Econometrics*, 11, 59–76.

Morris, A., Wilson, K., & Bakalis, S. (1995). Modeling tourism flows from Europe to Australia. *Tourism Economics*, 2, 147–167.

Porter-Hudak, S. (1990). An application of seasonal fractionally differenced model to the monetary aggregates. *Journal of the American Statistical Association*, 85, 338–344.

Ray, B. K. (1993). Long-range forecasting of IBM product revenues using a fractionally differenced ARMA model. *International Journal of Forecasting*, 9, 255–269.

Singapore Tourist Promotion Board. *Annual statistical report on visitor arrivals to Singapore—various issues 1975–2004*.

Song, H., & Witt, S. F. (2006). Forecasting international tourist flows to Macau. *Tourism Management*, 27, 214–224.

Sutcliffe, A. (1994). Time series forecasting using fractional differencing. *Journal of Forecasting*, 13, 183–393.

Turner, L. W., Kulendran, N., & Fernando, H. (1997). Univariate modeling using periodic and non-periodic analysis: Inbound tourism to Japan, Australia and New Zealand compared. *Tourism Economics*, 3, 39–56.

Turner, L. W., Kulendran, N., & Pergat, V. (1995). Forecasting New Zealand tourism demand with disaggregated data. *Tourism Economics*, 1, 51–69.

Uysal, M., & Crompton, J. (1985). An overview of approaches used to forecast tourism demand. *Journal of Travel Research*, 23, 7–15.

Van Doorn, J. W. M. (1984). Tourism forecasting and policy maker-criteria of usefulness. *Tourism Management*, 5, 24–39.

Witt, S. F., & Martin, C. (1987). Econometric models for forecasting international tourism demand. *Journal of Travel Research*, 25, 23–30.

Witt, C. A., Witt, S. F., & Wilson, N. (1994). Forecasting international tourist flows. *Annals of Tourism Research*, 21, 612–628.