농촌계획
Vol. 17, No. 1, 2011 (1-11)

# Estimating Monthly Tourist Population for Analysis of Green Tourism Potential in Village Level
## - A Case Study of Hahoe Village -

Gao, Yujie • Kim, Dae-Sik* • Kim, Yong-Hoon
*Department of Agricultural Engineering, Graduate School, Chungnam Nat'l Univv.*
*[*]Department of Agricultural Engineering, Chungnam Nat'l Univ.*

# 그린투어리즘 포텐셜 분석을 위한 관광마을 수준의 월별 방문객 추정
## - 하회마을을 중심으로 -

고옥결 • 김대식* • 김용훈
충남대학교 대학원 • *충남대학교 농업생명과학대학 지역환경토목학과

국문초록 : 본 연구에서는 ARIMA(Autoregressive Integrated Moving Average) 모델을 이용하여 농촌관광마을의 월별 관광객을 추정하였다. 단일 마을에 대한 시계열 자료를 경상북도 안동시에 위치한 하회마을을 대상으로 구축하였다. 월별 시계열 자료는 2000년부터 2010년까지 구성되었는데(2008년도 누락), 2000년에서 2007년까지 자료는 최적 모델의 도출에 나머지는 예측치의 검정에 사용되었다. 연구 결과 최적모델에 필요한 시계열 자료의 길이는 6년으로 나타났으며, 최적모델은 계절성을 고려한 SARIMA(2,1,1)(1,1,2)12로 나타났다. 최적 시계열 년수로 나타난 6년을 사용하여 2000-2005, 2001-2006, 그리고 2002-2007의 자료로부터 각각 SARIMA(2,1,1)(1,1,2)12를 도출하여, 차기년도들에 대한 예측결과를 비교한 결과, 높은 $R^2$값을 보였다.

**Key words** : ARIMA model; tourist arrival forecasting; tourist population; rural tourism.

## Introduction

Rural green tourism is becoming more and more popular because of its many characteristics, such as beautiful natural scenery, fresh air, traditional folk culture, simple and peaceful country life, and so on. Rural tourism can create more job opportunity, raise peasants' income, reduce out-migration, and diversify the rural economy. Seasonality is one of the most significant characteristics of tourism time series (Chang and Liao, 2010). So it is urgent to understand the future seasonal pattern of tourism in order to take effective measures in advance to (1) ensure the quality of tourism facilities and service and the sufficiency of general infrastructure for peak tourist season; (2) develop new tourism resources and products to promote off-season tourism.

ARIMA model has been widely used for tourism demand forecasting (Chen and Wang, 2007; Kim and Moosa, 2005; Cho, 2003; Chang and Liao, 2010; Huang and Min, 2002; Yim, 1990; Lim and McAleer, 2000; Cho, 2001; Goh and Law, 2002). Chang and Liao (2010) used seasonal ARIMA models to forecast the monthly outbound tourism departures of three major destinations from Taiwan to Hong Kong, Japan and the USA, respectively. Yim (1990) forecasted the tourism demand of American, Japanese, and Taiwanese visitors to Korea using ARIMA model and multiple regression model. Huang and Min

Corresponding author: Kim, Dae-Sik
Tel: 042-821-5795
E-mail: drkds19@cnu.ac.kr

(2002) predicted the number of visitor arrivals to Taiwan after the earthquake which happened in 1999 to evaluate whether Taiwan tourism had rebounded completely from the earthquake using SARIMA model. Lim and McAleer (2000) used ARIMA models to explain the nonstationary seasonally unadjusted quarterly tourist arrivals from Hong Kong and Singapore to Australia from 1975 to 1996. Cho (2001) applied three time-series forecasting techniques, namely, exponential smoothing, univariate ARIMA, and adjusted ARIMA to prediction of travel demand from different countries to Hong Kong, which showed that univariate ARIMA and adjusted ARIMA were more suitable and could be used to forecast the fluctuating series of visitor arrivals. Goh and Law (2002) presented the use of time series SARIMA and MARIMA with interventions in forecasting tourism demand using ten arrival series for Hong Kong. The forecasts obtained using models that capture stochastic nonstationary seasonality and interventions, SARIMA and MARIMA with intervention analysis were compared with other eight time series models and were found to have the highest accuracy. Although numerous studies have applied ARIMA model to prediction of the tourist arrivals, only few of them focused on rural tourist arrival forecasting (Son and Park, 2009), which they studied on quarterly prediction of tourists for farm stay villages using 5 years data of several hundreds' villages gathering from agents of cities and counties without a village-based data. And many studies have selected Hahoe Village for other purpose as a study area (Son and Park, 2009; Cho and Seo, 2009; Lee and Son, 2010; Jung, 2010; Hyun et al., 2005; Cho and Jang, 2005), but few of them forecasted its monthly tourist population with long time series data for a specific rural village. The monthly tourist population should be forecasted because tourism villages need to develop management and experience programs for the tourists seasonally and monthly.

The objective of this study is to establish a general model and determine the optimal length of time series for village-based monthly tourist arrival forecasting and pattern analysis. This study uses monthly time series data for 10 years of Hahoe Village, because the monthly tourists population and its pattern are so important for planning of village management programs such as experience program and stay plan in the village.

# Study area and data

## Description of study area

Hahoe Village is located in Andong, North Gyeongsang Province, Korea (Fig. 1). In Korean, this village is called Hahoe Maul which means river encircles village, because Nakdong River flows around the village in an S shape. The aristocratic Ryu clan's tile-roofed houses in the center of the village surrounded by farmers' thatched-roof ones preserve the architectural styles of the Joseon period. The village is also famous for the traditional Hahoe Byeolsingut Mask Dance, magnificent Buyongdae Cliff, sandy riverside, and ancient pine trees.

This village became more famous after the visit of Queen Elizabeth of England on April 21, 1999. Hahoe Village, along with Yangdong Village in Gyeongju, was added to the UNESCO World Heritage List under the category of “Historic Villages in Korea” on July 31, 2010 (Korea Tourism Organization).


Figure 1 Image view of HaHoe Village from Google Earth.

## Data Description

The monthly tourist arrival data from 2000 to 2010 except 2008 without data file were gotten from Hahoe Village Superintendent's Office. In this study, the tourist population refers to the tourists who paid for the entrance fees because of missing data of total tourist population. These monthly tourist arrival data are significant rare data for a village-based tourism in Korea. Because of lack of monthly data for other tourism villages, this study tried to get a general model that may not need mass data for monthly tourist arrival forecasting.

## Research methods

ARIMA model has three basic components: autoregression (AR), differencing or integration (I), and moving average (MA). The time series to be modeled should be a stationary series which has a constant mean and variance over time. ARIMA model building procedure consists of three steps: identification, estimation, and diagnosis. This procedure can be repeated until the best possible model is constructed for a series (SPSS Trends™ 13.0).

The general form of non-seasonal ARIMA model is ARIMA(*p,d,q*), where $p$ is the order of the non-seasonal autoregression, $d$ is the order of the non-seasonal differencing, and $q$ is the order of the non-seasonal moving average. These components are used to explain significant correlations found in the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots and to handle trends (SPSS Trends™ 13.0). The general form of the seasonal ARIMA model is SARIMA(*p,d,q*)(*P,D,Q*)s, where $P$, $D$, and $Q$ are the seasonal AR, I, and MA components (Koutroumanidisetal., 2009) and s is the seasonal periodicity (s=12 for monthly data). The ARIMA(*p,d,q*) model and SARIMA(*p,d,q*)(*P,D,Q*)s can be written as (Kim and Moosa, 2005):

ARIMA(*p,d,q*):

$$(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)(1-B)^d Y_t \\ = c+(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q)e_t \quad (1)$$

SARIMA(*p,d,q*)(*P,D,Q*)s:

$$(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)(1-\Phi_1 B^s-\Phi_2 B^{2s}-\cdots-\Phi_P B^{Ps})* \\ (1-B)^d(1-B^s)^D Y_t \\ = c+(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q) \\ (1-\Theta_1 B^s-\Theta_2 B^{2s}-\cdots-\Theta_Q B^{Qs})e_t \quad (2)$$

where $Y_t$ is a time series process; $e_t$ is a random error process; $c$ is a constant; $B$ is the backward shift operator; $\phi_1,\phi_2,\ldots\phi_p$ and $\Phi_1,\Phi_2,\ldots\Phi_P$ are the non-seasonal and seasonal autoregressive coefficients, respectively; $\theta_1,\theta_2,\ldots\theta_q$ and $\Theta_1,\Theta_2,\ldots\Theta_Q$ are the non-seasonal and seasonal moving average coefficients, respectively.

This study used ARIMA model to get a general model and determine the optimal length of time series for village-based monthly tourist arrival forecasting and pattern analysis which are very important for planners to make timely and resonable decisions for green tourism development. Because of lack of data for 2008, data from 2000 to 2007 were used for model calibration. In order to find a general model which can be obtained using data as short a time period as possible, 8-year, 7-year, 6-year, and 5-year time series were used for model building, respectively. The forecast accuracy measures used are the R-squared ($R^2$), root mean square error (RMSE), mean absolute percentage error (MAPE), and Normalized Bayesian Information Criteria (Normalized BIC). 10 alternative models were selected for each length of time series. After comparison of the alternative models, the general model was gotten and then verified using three 6-year time series for model building and the rest of years' data for model verification.

## Results and Discussion

### General model selection

The identification process requires a stationary time series. The results show that although data from different time series were used for model calibration, they all need first-order non-seasonal differencing and first-order seasonal differencing to make the series stationary, taking the time series 2002-2007 for example.

Fig. 2 shows the monthly tourist population of Hahoe Village from 2000 to 2007. The series shows a downward trend as a whole, making it clear that the series is not stationary. The time series also has many peaks, many of which appear to be equally spaced. This means that the time series has the peak tourist season (April, May, August, and October). Fig. 3 shows the series differencing which can make a series stationary. Fig. 3 (a) is the ACF of time series and it shows that the autocorrelations at lags 12 and 24 are unusually high, which confirms the presence of the seasonal component. The first-order non-seasonal differencing was first performed as shown in Fig. 3 (b)-(c).

The time series plot exhibits that the series seems to be stationary but the small drop in the ACF at lag 24 relative to lag 12 indicates that the series is not stationary and the seasonal differencing is necessary. After first-order non-seasonal differencing and first-order seasonal differencing, the time series has been stabilized as shown in Fig. 3 (d)-(f). Now determining any autoregressive and/or moving-average orders is needed to model the series.

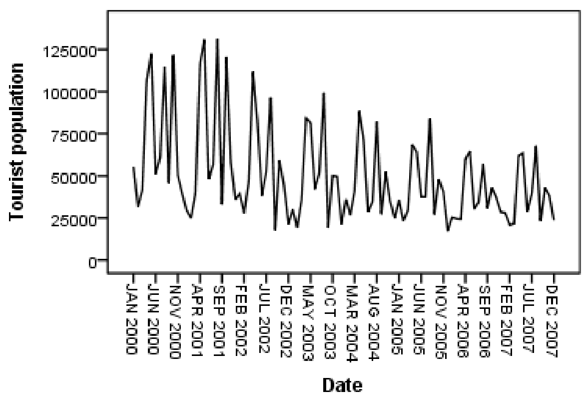


Figure 2 Monthly tourist population from 2000 to 2007.

Table 1 shows the alternative models obtained using data from different lengths of time series. The models with high $R^2$ were first considered because the general model obtained will be used for monthly tourist arrival pattern analysis. 10 models were obtained for each series. The comparison results show that SARIMA(2,1,1)(1,1,2)12 model can be the general model because the model exists in all the lengths of time series. Another result is that the highest $R^2$ value is 0.871 which is obtained by using 6-year time series. This indicates that using 6-year series is

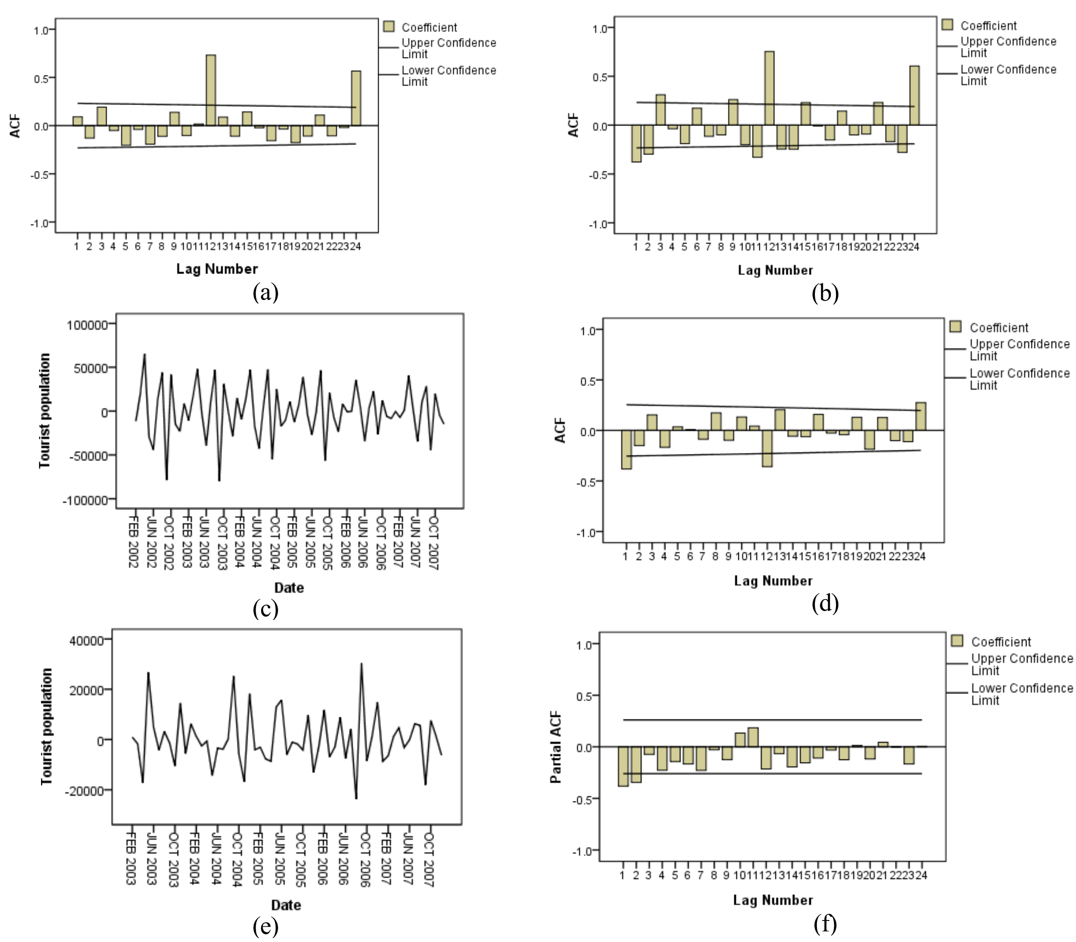


Figure 3 Time series differencing for time series 2002-2007.
(a) ACF of time series; (b)-(c) Non-seasonally differenced ACF and plot; (d)-(f) Non-seasonally and seasonally differenced ACF, plot, and PACF.

Table 1 Comparison of the models for general model selection.

| Lengths of time series (year) | Time series | Models | Model fit statistics | | | |
|---|---|---|---|---|---|---|
| | | | $R^2$ | RMSE | MAPE | Normalized BIC |
| 8 | 2000-2007 | SARIMA(2,1,1)(2,1,0)12 | 0.818 | 12034.185 | 18.142 | 19.110 |
| | | SARIMA(2,1,1)(1,1,1)12 | 0.818 | 12035.691 | 18.275 | 19.111 |
| | | SARIMA(1,1,2)(1,1,1)12 | 0.818 | 12036.242 | 18.300 | 19.111 |
| | | SARIMA(1,1,2)(1,1,0)12 | 0.817 | 11968.052 | 18.313 | 19.046 |
| | | SARIMA(2,1,1)(0,1,2)12 | 0.817 | 12068.415 | 18.298 | 19.116 |
| | | SARIMA(2,1,1)(1,1,0)12 | 0.816 | 12002.251 | 18.114 | 19.052 |
| | | SARIMA(2,1,1)(1,1,2)12 | 0.818 | 12115.905 | 18.153 | 19.177 |
| | | SARIMA(2,1,1)(2,1,1)12 | 0.818 | 12114.983 | 18.153 | 19.177 |
| | | SARIMA(2,1,1)(2,1,2)12 | 0.817 | 12202.597 | 18.281 | 19.245 |
| | | SARIMA(2,1,2)(0,1,2)12 | 0.816 | 12155.032 | 18.306 | 19.184 |
| 7 | 2001-2007 | SARIMA(2,1,1)(0,1,2)12 | 0.708 | 12594.583 | 19.401 | 19.242 |
| | | SARIMA(2,1,1)(1,1,2)12 | 0.708 | 12691.551 | 19.489 | 19.318 |
| | | SARIMA(2,1,1)(2,1,1)12 | 0.708 | 12697.240 | 19.372 | 19.319 |
| | | SARIMA(1,1,2)(1,1,1)12 | 0.706 | 12645.034 | 19.351 | 19.250 |
| | | SARIMA(1,1,2)(2,1,1)12 | 0.707 | 12717.312 | 19.252 | 19.322 |
| | | SARIMA(2,1,1)(2,1,0)12 | 0.707 | 12607.101 | 19.197 | 19.244 |
| | | SARIMA(2,1,1)(2,1,2)12 | 0.707 | 12826.489 | 19.437 | 19.399 |
| | | SARIMA(2,1,2)(2,1,1)12 | 0.706 | 12832.863 | 19.271 | 19.400 |
| | | SARIMA(2,1,2)(2,1,2)12 | 0.706 | 12939.854 | 19.368 | 19.476 |
| | | SARIMA(2,1,1)(0,1,1)12 | 0.704 | 12584.572 | 19.618 | 19.181 |
| 6 | 2002-2007 | SARIMA(1,1,2)(1,1,2)12 | 0.871 | 7798.735 | 15.683 | 18.407 |
| | | SARIMA(2,1,2)(1,1,2)12 | 0.871 | 7870.331 | 15.703 | 18.495 |
| | | SARIMA(1,1,2)(0,1,2)12 | 0.870 | 7740.257 | 15.525 | 18.323 |
| | | SARIMA(1,1,2)(2,1,2)12 | 0.870 | 7908.129 | 15.536 | 18.504 |
| | | SARIMA(2,1,1)(1,1,2)12 | 0.870 | 7826.876 | 15.792 | 18.414 |
| | | SARIMA(2,1,2)(0,1,2)12 | 0.870 | 7817.448 | 15.631 | 18.412 |
| | | SARIMA(1,1,1)(1,1,2)12 | 0.869 | 7767.997 | 15.797 | 18.330 |
| | | SARIMA(2,1,1)(0,1,2)12 | 0.869 | 7787.504 | 15.569 | 18.335 |
| | | SARIMA(2,1,1)(2,1,2)12 | 0.869 | 7914.835 | 15.434 | 18.506 |
| | | SARIMA(2,1,2)(2,1,2)12 | 0.869 | 8008.462 | 15.699 | 18.599 |
| 5 | 2003-2007 | SARIMA(1,1,2)(2,1,1)12 | 0.825 | 8405.852 | 16.486 | 18.647 |
| | | SARIMA(0,1,2)(1,1,1)12 | 0.824 | 8226.593 | 16.397 | 18.440 |
| | | SARIMA(1,1,2)(2,1,0)12 | 0.824 | 8315.375 | 16.261 | 18.543 |
| | | SARIMA(2,1,1)(1,1,1)12 | 0.824 | 8333.912 | 16.776 | 18.548 |
| | | SARIMA(2,1,1)(1,1,2)12 | 0.824 | 8431.720 | 16.641 | 18.653 |
| | | SARIMA(2,1,1)(2,1,1)12 | 0.824 | 8434.143 | 16.720 | 18.654 |
| | | SARIMA(2,1,2)(1,1,1)12 | 0.824 | 8413.907 | 16.494 | 18.649 |
| | | SARIMA(2,1,2)(2,1,1)12 | 0.824 | 8527.178 | 16.498 | 18.757 |
| | | SARIMA(0,1,2)(1,1,2)12 | 0.824 | 8325.163 | 16.553 | 18.546 |
| | | SARIMA(0,1,2)(2,1,1)12 | 0.824 | 8325.758 | 16.633 | 18.546 |

the most appropriate for model calibration. Then SARIMA(2,1,1)(1,1,2)12 model as a general model for village-based monthly tourist arrival forecasting was verified using three 6-year time series (2000 to 2005, 2001 to 2006, and 2002 to 2007) for model building and the rest of years' data for model verification.

Figure 3 Time series differencing for time series 2002-2007. (a) ACF of time series; (b)-(c) Non-seasonally differenced ACF and plot; (d)-(f) Non-seasonally and seasonally differenced ACF, plot, and PACF.

## General model verification

Model verification is usually required when ARIMA models are applied to forecasting time series. In this study, the proposed SARIMA(2,1,1)(1,1,2)12 model was assessed by comparing the observed and predicted data. Three 6-year time series 2000-2005, 2001-2006, and 2002-2007 were used for model building and the rest of years' data were used for model verification, respectively.

### Model 1: Model building using time series 2000-2005

Fig. 2 and Fig. 4 (a) show that the time series 2000-2005 was unstationary just like the series 2002-2007. In order to verify the SARIMA(2,1,1)(1,1,2)12 model using series 2000-2005, the first step is to see if the series is stationary after first-order non-seasonal differencing and first-order seasonal differencing. The results shown in Fig. 4 (b)-(d) exhibit that the series has been stabilized after first-order non-seasonal and seasonal differencing. This means that the SARIMA(2,1,1)(1,1,2)12 model can be applied to the series. The statistic data of SARIMA(2,1,1)(1,1,2)12 model are shown in Table 2 and the model equation is Eq. (3).

The SARIMA(2,1,1)(1,1,2)12 model equations obtained using three 6-year time series can be written as:

Model 1:

$$(1-0.224B-0.123B^2)(1+0.432B^{12})(1-B)(1-B^{12})Y_t = 96.581+(1-B)(1+0.208B^{12}+0.115B^{24})e_t \quad (3)$$

Model 2:

$$(1-0.191B-0.121B^2)(1-0.371B^{12})(1-B)(1-B^{12})Y_t = 157.494+(1-0.997B)(1-0.599B^{12}+0.379B^{24})e_t \quad (4)$$

Model 3:

$$(1-0.195B+0.053B^2)(1+0.371B^{12})(1-B)(1-B^{12})Y_t = 27.778+(1-0.997B)(1+0.043B^{12}+0.995B^{24})e_t \quad (5)$$

Diagnosing an ARIMA model involves analyzing the model residuals. Residuals of the model shown in Fig. 5 are random and all the lags are within 95% confidence interval. This means that the model is satisfactory. The comparison results between observed and predicted data shown in Figures 6-9 exhibit high goodness of fit with $R^2$ of 0.85, 0.93, 0.90, and 0.82 for 2006, 2007, 2009, and 2010, respectively.

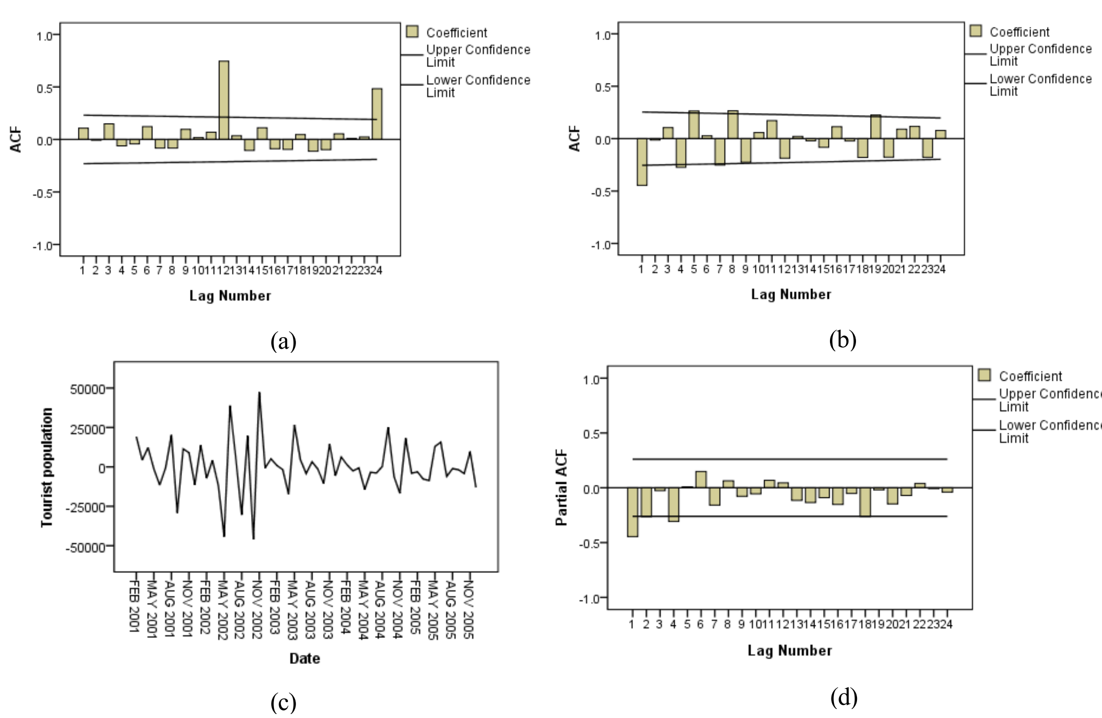


Figure 4 Time series differencing for time series 2000-2005.
(a) ACF of time series; (b)-(d) Non-seasonally and seasonally differenced ACF, plot, and PACF.

Table 2 SARIMA(2,1,1)(1,1,2) model statistics obtained using three 6-year time series.

| Items | | Year | Time series models | | |
|---|---|---|---|---|---|
| | | | Model 1 (2000-2005) | Model 2 (2001-2006) | Model 3 (2002-2007) |
| Model building | Model Parameters | Constant | 96.581 | 157.494 | 27.778 |
| | | AR(1) | 0.224 | 0.191 | 0.195 |
| | | AR(2) | 0.123 | 0.121 | -0.053 |
| | | MA(1) | 1.000 | 0.997 | 0.997 |
| | | SAR(12) | -0.432 | 0.371 | -0.371 |
| | | SMA(12) | -0.208 | 0.599 | -0.043 |
| | | SMA(24) | -0.115 | -0.379 | -0.995 |
| | Model Fit statistics | $R^2$ | 0.808 | 0.683 | 0.870 |
| | | RMSE | 13877.701 | 13835.643 | 7826.876 |
| | | MAPE | 20.070 | 20.885 | 15.792 |
| | | Normalized BIC | 19.560 | 19.554 | 18.414 |
| Verification | $R^2$ with data year | 2006 | 0.85 | - | - |
| | | 2007 | 0.93 | 0.94 | - |
| | | 2009 | 0.90 | 0.75 | 0.80 |
| | | 2010 | 0.82 | 0.64 | 0.61 |

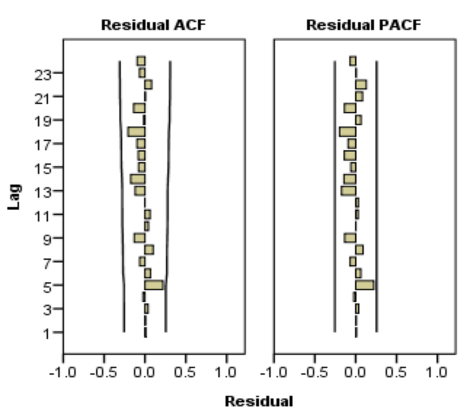


Figure 5 Residuals of Model 1.

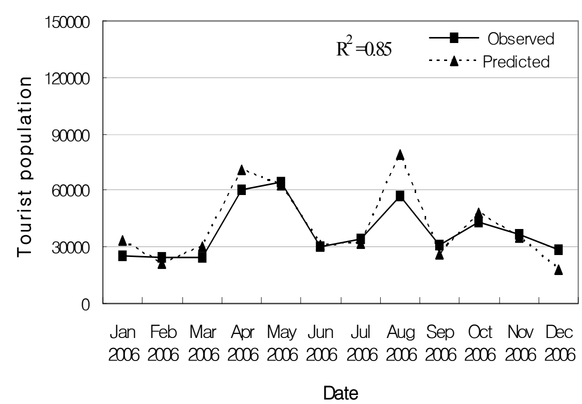


Figure 6 Prediction result for 2006 using Model 1.

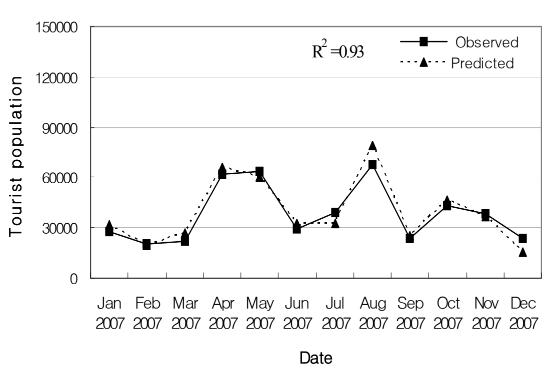


Figure 7 Prediction result for 2007 using Model 1.

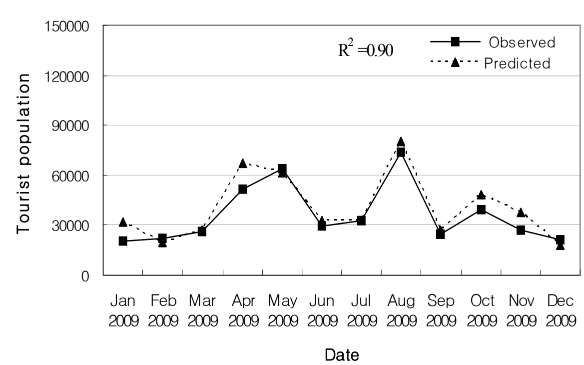


Figure 8 Prediction result for 2009 using Model 1.

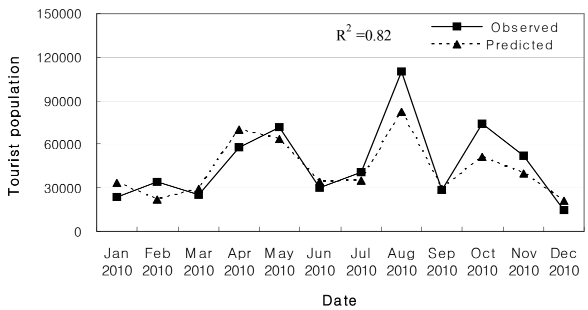


Figure 9 Prediction result for 2010 using Model 1.

differencing. It indicates that SARIMA(2,1,1)(1,1,2)12 model can be applied to the time series. The statistic data are shown in Table 2 and the model equation can be written as Eq. (4).

Residuals of the model shown in Fig. 11 are random and all the lags are within 95% confidence interval. This means that the model is adequate. The comparison results between the observed and predicted data (Figures 12-14) show high goodness of fit with $R^2$ of 0.94, 0.75, and 0.64 for 2007, 2009, and 2010, respectively.

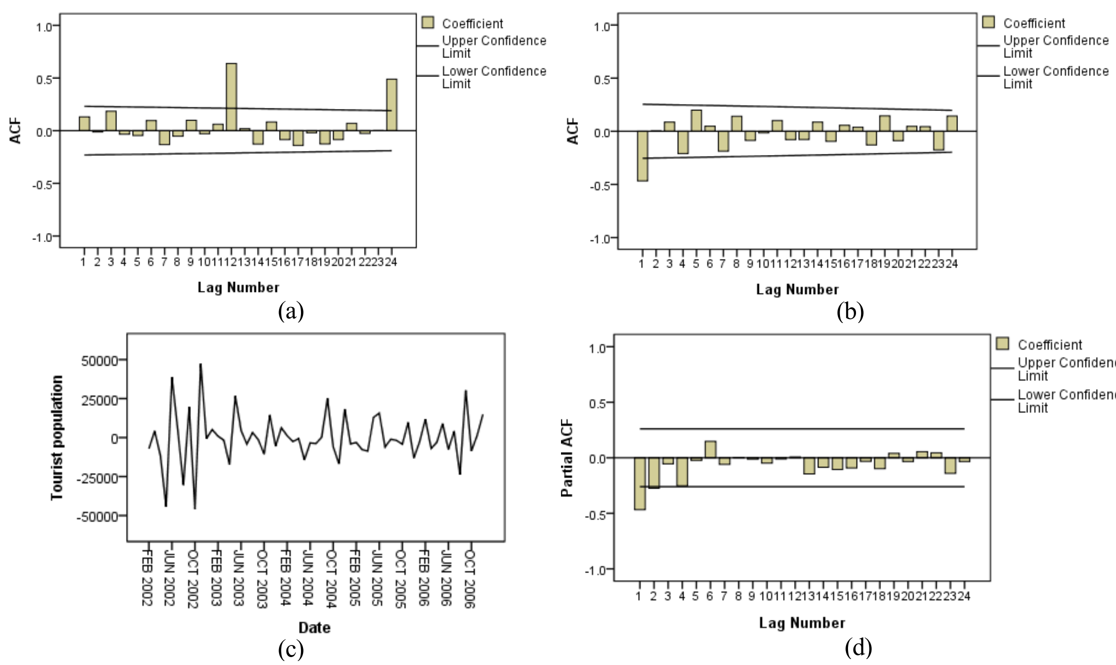


Figure 10 Time series differencing for time series 2001-2006.
(a) ACF of time series; (b)-(d) Non-seasonally and seasonally differenced ACF, plot, and PACF.

### Model 2: Model building using time series 2001–2006

Here, the time series 2001-2006 was used for model building and data for 2007, 2009, and 2010 were used for model verification. The time sequence plot and ACF shown in Fig. 2 and Fig. 10 (a) reveal that the time series is unstationary. In order to verify the SARIMA(2,1,1)(1,1,2)12 model, this time series was also stabilized using the first-order non-seasonal differencing and first-order seasonal differencing. The ACF, time series plot, and PACF shown in Fig. 10 (b)-(d) show that the time series is stationary after first-order non-seasonal and first-order seasonal

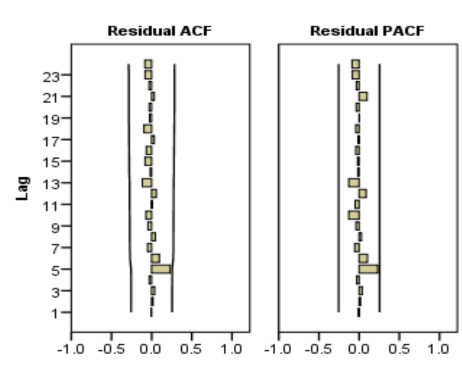


Figure 11 Residuals of Model 2.

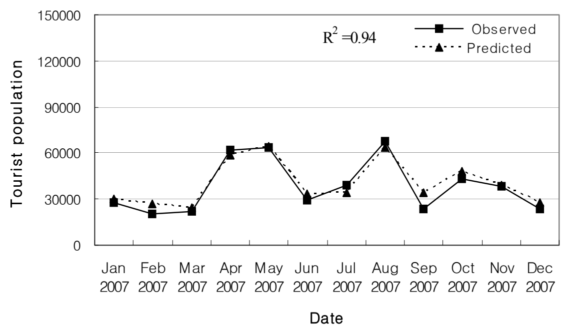

Figure 12 Prediction result for 2007 using Model 2.

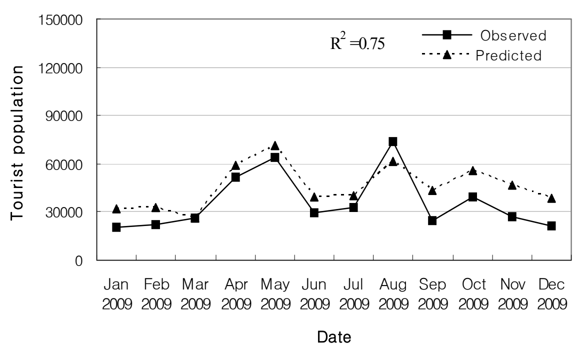

Figure 13 Prediction result for 2009 using Model 2.

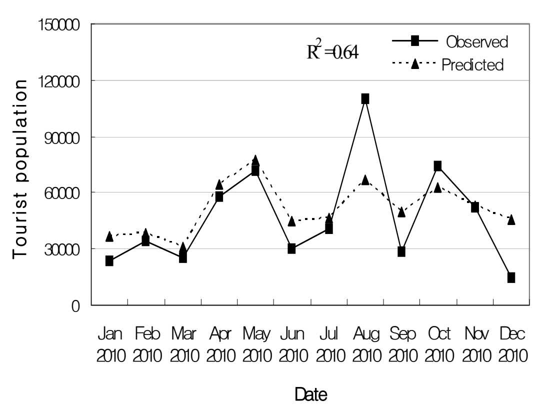

Figure 14 Prediction result for 2010 using Model 2.

## Model 3: Model building using time series 2002–2007

This time series 2000-2007 were used for model-building and data from 2009 to 2010 were used for model verification. The series has been stabilized using first-order non-seasonal differencing and first-order seasonal differencing as shown in Fig. 3 (d)-(f). The equation of the SARIMA(2,1,1)(1,1,2)12 model can be written as Eq. (5).

Residuals of SARIMA(2,1,1)(1,1,2)12 model shown in Fig. 15 for diagnosing the model are random and all the lags are within 95% confidence interval, which reveals that the model is satisfactory. And then the model was verified using the data from 2009 to 2010. The comparison results between observed and forecasted data were shown in Figures 15-16 with $R^2$ of 0.80 for 2009 and 0.61 for 2010.

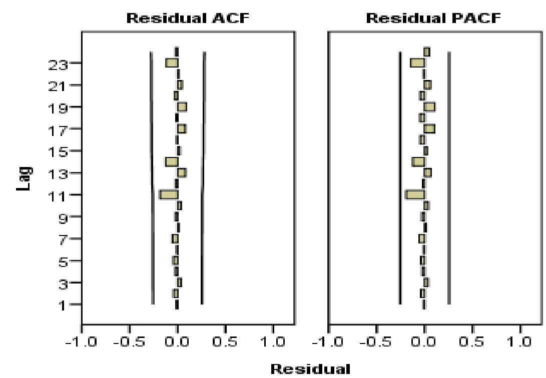

Figure 15 Residuals of Model 3.

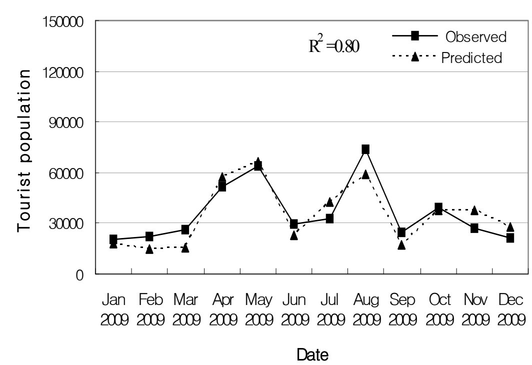

Figure 16 Prediction result for 2009 using Model 3.

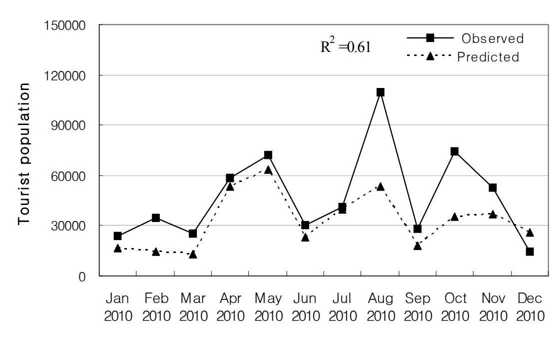

Figure 17 Prediction result for 2010 using Model 3.
Summary and conclusion.

## Summary and Conclusion

In this study, ARIMA forecasting method which has not been widely applied to rural green tourism was used to build a general model and determine the optimal length of time series for village-based monthly tourist arrival forecasting and pattern analysis. The data gotten from Hahoe Village Superintendent's Office were monthly tourist arrival data that are significant rare data for tourism villages in Korea. Using four different lengths of time series and ten alternative models for each time series, this study found that 6-year time series with the highest $R^2$ values is the best length of time series and SARIMA(2,1,1)(1,1,2)12 model which exists in all the lengths of time series was selected as the general model to forecast village-based tourist arrivals. The proposed general model was verified using three 6-year time series for model building and the rest of years' data for model verification. The results show that most of $R^2$ values between observed and predicted data are more than 0.7 except two. This means that SARIMA(2,1,1)(1,1,2)12 model and 6-year data are adequate for village-based monthly tourist arrival forecasting.

Although the number of tourist arrivals can be easily affected by social and economical factors, forecasting tourist arrivals is still necessary. The results of this study are very useful for local governments of village-based tourism attractions, expecially those that lack monthly tourist arrival data, to make timely and reasonable tourism policy. On the one hand, it’s necessary to improve the quality and range of general infrastructure, tourism facility and service for peak tourist season. On the other hand, it’s essential to develop new tourism resources and products to promote the local off-season tourism. What is more important is sustainable tourism which needs the conservation of local environment and natural and cultural heritage.

## References


1. Chang, Y. W. and M. Y. Liao, 2010, A seasonal ARIMA model of tourism forecasting: The Case of Taiwan, Asia Pacific Journal of Tourism Research, 15(2), 215-221.
2. Che, J. and J. Wang, 2010, Short-term electricity prices forecasting based on support vector regression and Auto-regressive integrated moving average modeling, Energy Conversion and Management, 51, 1911-1917.
3. Chen, K. Y. and C. H. Wang, 2007, Support vector regression with genetic algorithms in forecasting tourism demand, Tourism Management, 28, 215-226.
4. Cho, M. H. and H. J. Jang, 2005, A study on residents' attitudes to the utilization of traditional culture as tourism resources: Based on Jirisan Cheonghak-dong and Hahoi Village, Journal of Tourism and Leisure Research, 17(1), 133-154 (in Korean).
5. Cho, T. Y. and T. Y. Seo, 2009, Effect of Cultural Tour Experience Hahoe Village at Andong on Touring Satisfaction and Behavioral Intention, The Journal of the Korea Contents Association, 9(7), 361-370 (in Korean).
6. Cho, V., 2001, Tourism forecasting and its relationship with leading economic indicators, Journal of Hospitality and Tourism Research, 25, 399-420.
7. Cho, V., 2003, A comparison of three different approaches to tourist arrival forecasting, Tourism Management, 24, 323-330.
8. Gil-Alana, L. A., 2005, Modelling international monthly arrivals using seasonal univariate long-memory processes, Tourism Management, 26, 867-878.
9. Goh, C. and R. Law, 2002, Modeling and tourist arrival forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention, Tourism Management, 23, 499-510.
10. Huang, J. H. and J. C. H. Min, 2002, Earthquake devastation and recovery in tourism: The Taiwan case, Tourism Management, 23, 145-154.
11. Hyun, Y. H., S. H. Han, and H. J. Huh, 2005, Relationship between the destination image formation and tourist behaviour based on familiarity index: A study of Hahoe Village, International Journal of Tourism Sciences, 29(1), 147-167 (in Korean).
12. Jia, J. S., J. Z. Zhao, H. B. Deng, and J. Duan, 2010, Ecological footprint simulation and prediction by ARIMA model-A case study in Henan Province of

China, Ecological Indicators, 10, 538-544.
13. Jung, Y. S., 2010, A Study on the location of house and geographical features of Hahoe Village, Journal of the Architectural Institute of Korea Planning & Design, 26(7), 135-144 (in Korean).
14. Kim, J. H. and I. A. Moosa, 2005, Forecasting international tourist flows to Australia: a comparison between the direct and indirect methods, Tourism Management, 26, 69-78.
15. Korea Tourism Organization: Andong Hahoe Village, http://www.visitkorea.or.kr/enu/SI/SI_EN_3_1_1_1.jsp?cid=264148.
16. Koutroumanidis, T., K. Ioannou, and G. Arabatzis, 2009, Predicting fuelwood prices in Greece with the use of ARIMA models, artificial neural networks and a hybrid ARIMA–ANN model, Energy Policy, 37, 3627-3634.
17. Lee, H. S. and Y. H. Son, 2010, A study on landscape changes in a riverside village due to floods and public works for flood control: Focused on the Nakdong River Hahoe Village in Korea, Journal of the Korean Planners Association, 45(3), 243-252 (in Korean).
18. Lim, C. and M. McAleer, 2000, A seasonal analysis of Asian tourist arrivals to Australia, Applied Economics, 32, 499-509.
19. Son, E. H. and D. B. Park, 2009, A forecasting of farm stay tourists on rural tourism villages, Green Tourism Research, 16(4), 27-44 (in Korean).
20. SPSS Trends™ 13.0, 2004, SPSS Inc.
21. Valenzuela, O., I. Rojas, F. Rojas, H. Pomares, L. J. Herrera, A. Guillen, L. Marquez, and M. Pasadas, 2008, Hybridization of intelligent techniques and ARIMA models for time series prediction, Fuzzy Sets and Systems, 159, 821-845.
22. Yim, E. S., 1990, A study on the demand forecasting model of American, Japanese and Taiwanese visitors to South Korea, Study on Tourism, 14, 141-156 (in Korean).

접 수 일: (2010년 8월 9일)
수 정 일: (1차: 2010년 9월 16일, 2차: 2011년 2월 8일)
게재확정일: (2011년 2월 8일)
■ 3인 익명 심사필