Forecasting Tourist Inflow in Bhutan using Seasonal ARIMA

Elangbam Haridev Singh[1]

Senior Lecturer, Gaeddu College of Business Studies, Royal University of Bhutan, Bhutan

Abstract: ***The main aim of this paper is to generate one-period-ahead forecasts of international tourism demand for Bhutan by selecting appropriate model both ARIMA as well as exponential smoothing. Before selecting an appropriate model, formal stationary tests has been applied in this paper and finds that, the series are stationary at level. Secondly, in order to get a good estimation, this paper has identified the autoregressive (AR) and moving average (MA) of the entire period of the data. Therefore, the future demand of tourism is forecast based on the combination of AR and MA, which known as ARMA model. In this paper, the competing models have been thoroughly investigated when the model adequacy has been checked before the best combination of ARIMA model was selected. During the diagnostic stage we came to know that the data has a seasonal components thus, the best fitted ARIMA (0, 1, 1) (1, 1, 1) with seasonal effects or well known as SARIMA approaches has been suggested through this study and the forecasting process is based on this combination.***

Keywords: Forecasting, Tourism, ARIMA, Seasonal ARIMA.

1. Introduction

Tourism industry is a vital source of economic growth and development in many countries around the world but the case is more prominent in developing countries like Maldives (42% of the GDP was contributed by Tourism in 2006), Azerbaijan(20.33% of GDP), Cambodia (13.6% of GDP (2006)), Hong Kong, China (10.27 % of GDP (2006)), Philippines (16.37 % of GDP (2006)), Bhutan (6% of GDP in 2009 second only to Hydropower), etc. The Travel & Tourism Economy had generate over 238 million jobs worldwide and contribute to 9.9% of global GDP in 2008, expected to provide more than 296 million jobs and 10.5% global GDP by 2018. International tourist arrivals and international tourist receipts have usually been used as benchmark aggregate series to assess the overall importance of tourism worldwide and in specific countries. High international tourist arrival levels signify peace and prosperity of a country in the international community. Also, sizeable international tourist receipts can be a good indicator of the role of tourism in an economy in terms of both Gross Domestic Product and foreign exchange generation in developing countries and LDCs.

During the period 1995-1998, tourism revenues were one of the five leading sources of forex earning export revenue for 69 developing countries. Among the latter, tourism revenue was the main source of foreign currency in 28 countries, its share in total exports ranging between 79 and 20 per cent; in 27 countries it accounted for between 20 and 10 per cent; and in the 24 remaining countries it was around 10 per cent. The contribution of export revenues to gross domestic product (GDP) was highly important and accounted for 82.29 per cent (in Maldives) and 30 per cent (in Samoa). One aspect which is evident from UN (2007) study is that although the contribution of tourism revenues is significant in all the countries, its contribution to GDP is declining as the economies become more diversified. Mauritius, the Dominican Republic, Bhutan and Tunisia are the perfect example.

For Bhutan 2010 was a significant year as the number of tourist arrival in the country is beyond the expectation of the industry estimate. The country, offer hospitality to 40,873 high-end tourists in 2010 exceeding its own target of 35,000 tourists for the year which was nearly 17% more and earned more than Nu 539.21 million as revenue. The addition of job by this sector is also significant as new 2600 jobs were added and is estimate that between 17800 and 19600 (direct and indirect) people are employed by this sector so far. This figure gives an optimistic view and well exceeds the industry estimated of 17000 jobs. This research paper aimed is use to predict the number of international tourist arrivals to Bhutan in the period 2013 and 2014, and to seek provide the best model for forecasting international tourist arrivals to Bhutan for these periods using Box Jenkins Methods with seasonal modification.. For the study, we used the monthly data of number of tourist arrive in Bhutan from 1983-2012.

2. A Brief Survey of Literature

The literature concerning the tourist forecasting is abundantly available with various type of empirical analysis in both cross-sectional as well as time series data. But most of forecasting tourist demand used pure time-series analytical models and only a few are available on cross sectional analysis. Frequently encounter the time series forecasting methods are Holt's Additive and multiplicative smoothing methods, ARIMA, VAR, GRACH etc. Among these ARIMA using both seasonal and non seasonal is very common. Basically, this ARIMA model has been proved to be reliable in modeling and forecasting tourism demand with monthly and quarterly time-series. Wong has used four types of model, such as 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 [11]. The empirical results shows that forecast combinations do not always outperform the best single forecasts. Therefore, combination of empirical models can reduce the risk of forecasting failure in practical. Coshall used univariate analysis, combining the ARIMA-volatility and smoothing model, finance to forecast United Kingdom demand for international tourism [5]. Generally, from this study we can find that the ARIMA volatility models tend to overestimate demand, and the smoothing models are inclined underestimate the number of future tourist arrivals. Chu has modified ARIMA modeling to fractionally integrated autoregressive moving average (ARFIMA) in forecasting tourism demand [3]. This ARFIMA model is ARMA based methods. Three types of univariate models have applied in the study with some modification in ARMA model become ARAR and ARFIMA model. The main purpose of the 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 various forecasting tourism study which been publish earlier, because we can identify the ARMA based model behaviors and the outperforming of the ARFIMA model with other ARMA based models. Again, Chu has study the ARIMA based model using ARAR algorithm model in order to analyze and forecasting tourism demand for Asia-Pacific region using monthly and quarterly data [4]. The major findings of the studies show the comparison between forecasts generated by monthly and quarterly data reveals that the performance is broadly similar. Besides forecasting tourist arrivals, prediction of tourism revenue also can done using empirical modeling. Akal, M. has used autoregressive integrated moving average cause-effect (ARMAX) modeling 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 strength-out the tourism sector becoming a major contributing sector for economic development in future [1]. In the case of Bhutan there is no literature attempt to forecast the tourism demand. So, this paper can be a first attempt in this line.

3. Methodology

3.1 ARIMA Modeling

ARIMA models are the most general class of models for time series forecasting which can be stationarized by transformations such as differencing and logging. It was introduced by Box and Jenkins (1970) which includes autoregressive as well as moving average parameters, and explicitly includes differencing in the formulation of the model [1]. The figure 1 below gives the detail flow chart of carrying out ARIMA modeling.

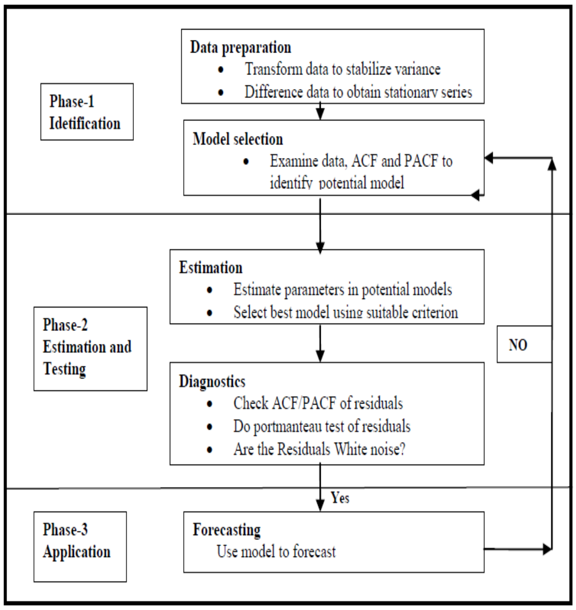

Figure 1: Diagramatic representation of forecasting using Bos-Jenkins models)

3.2 Description about Seasonal ARIMA:

Since the usual arima model ARIMA (p, d, q) can't handle the data, due to its seasonal component, we use ARIMA (p, d, q) (P, D, Q)s **).** Seasonal ARIMA models are defined by 7 parameters ARIMA (p, d, q) (P, D, Q)s and the equation I below model of seasonal ARIMA.

$$(1-\phi_1 B-\phi_2 B^2-..\phi_p B^p)(1-\beta_1 B^s-..-\beta_P B^{ps})(1-B)^d(1-B^s)^D y_t = \quad \text{Eq. 1}$$
$$c+(1-\psi_1 B-\psi_2 B-......-\psi_q B^q)(1-\theta_1 B-\theta_2 B^{2s}-....-\theta_Q B^{Qs})\varepsilon_t$$

where

- AR(p) Autoregressive part of order p
- MA(q)Moving average part of order q
- I (d) differencing of order d
- ARs (P) Seasonal Autoregressive part of order P
- MAs (Q) Seasonal Moving average part of order Q
- Is (D) seasonal differencing of order D
- s is the period of the seasonal pattern appearing i.e.
- = 12 months in the tourist inflow data.

Characteristics of a good ARIMA model are that (a) it is parsimonious, i.e. the model uses a small number of parameters or coefficients to explain the available data, (b) it is stationary, (c) it is mvertible, (d) it has high quality estimated coefficients, (e) it has uncorrelated residuals, (f) it fits the available data satisfactorily, that is, it has acceptable root-mean-squared error (RMSE) and mean absolute percent error (MAPE) values, and (g) it produces accurate forecasts. Advantages of ARIMA models are: (a) ARIMA models are based on a solid foundation of classical probability theory and mathematical statistics, (b) Versatility of the ARIMA models for various types of time series. For example, exponential smoothing methods can handle only time series with deterministic trend in which the average level of the series grows or declines according to some specific function in time. However, many time series displays stochastic trends, i.e. tends with variable rates of growth. Both deterministic and/or stochastic trend can be modeled using

ARIMA models, and (c) Unified, theoretically sound approach to forecasting. That is, the Box-Jenkins methodology provides a systematic approach to model selection, utilizing all the information available in the sample acf and pacf. Disadvantages of the Box-Jenkins method are: (a) the method is fairly complicated and considerable expertise is required in specifying a suitable model using the sample and partial autocorrelation functions, particularly in small samples, (b) difficulties of interpretation, and (c) attempts to select ARIMA models by an automatic procedure, such as the AIC can lead to worse results. Consequently, statistical packages which purport to offer automatic ARIMA selection are severely limited, (d) the method requires about at least 50 observations as the minimum amount of data in the time series, and (e) it does not necessarily guarantee better forecasts than simpler forecasting techniques.

3.3 Data preparation and model selection

The input series for ARIMA needs to be stationary. A stationary series should have a constant mean, variance, and autocorrelation through time. The purpose of identification phase is to determine the differencing required for producing stationary and also the order of non seasonal AR and MA operators for a given series. When the observed time series presents trend, differencing and transformation are often applied to the data to remove the trend and stabilize variance before an ARIMA model can be fitted. The input time series show upward trend with spike see. Figure no2. Then we differentiated onetime non seasonal and one time seasonal, the result being stabilized over mean (see figure no 3). But and we stabilized over variance by natural log transformation see figure no. 4)

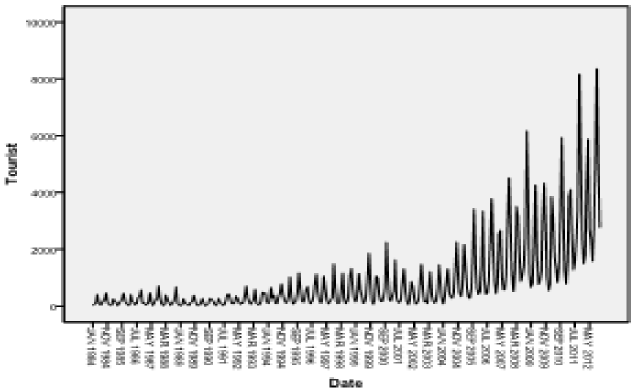

Figure 2: Source: Spss 17

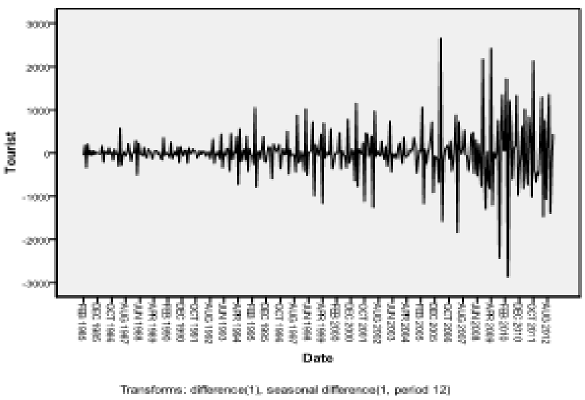

Figure 3: Source: Spss 17

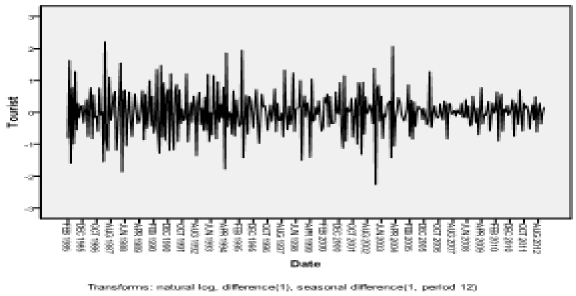

Figure 4: Source: Spss17

A precise correspondence between ARMA (p, q) processes and defined ACF and PACF repeated several times to obtain satisfactory model. This model is then used for prediction purposes. Patterns are more difficult to recognize. When the order of at least one of the two components (AR or MA) is clearly detectable, the other can be identified by attempts in the following step of parameter estimation. Finally, the existence of a seasonal component of length s is underlined by the presence of a periodic pattern of period s in the ACF.

3.4. Phase II (Estimation and Testing)

After identifying the suitable ARIMA (p, d, q) (P, D, Q) s structure, subsequent steps of parameter estimation and testing are performed. Estimation stage consists of using the data to estimate and make inferences about parameters of tentatively identified model. The parameters are estimated such that an overall measure of residuals is minimized. The last stage of model building is the testing or diagnostic checking of model adequacy. This stage determines whether residuals are independent, homoscedastic and normally distributed. Several diagnostic statistics and plots of the residuals are used to examine the goodness of fit. After identifying tentative model, the process is again followed by the stage of parameter estimation and model verification. Diagnostic information may help to suggest alternative model(s). Now, the series is stationary and several models were selected base on their ability of reliability predication. We are considering the AR, MA and the seasonal as well as seasonal aspects of the models, and estimates base on the criterions of BIC, R-RMSE, , R-Square and the Mean Absolute percentage were used. We use lower value of BIC, RMSE, MAPE and High value of R-square (see table below).

Table 1: Model Summary

Models	*R2*	*RMSE*	*MAPE*	*BIC*
ARIMA(0,1,1) (1,1,1)	.894	418.06	30.191	12.12
ARIMA(0,1,1) (2,1,1)	.905	299.127	28.667	11.531
ARIMA(0,1,1) (2,2,1)	.884	333.996	39.243	12.737
ARIMA(0,1,1)(2,2,2)	.892	323.046	39.95	12.689
ARIMA(0,1,1) (1,1,3)	.884	330.940	30.78	12.733
ARIMA(0,1,1) (2,1,2)	.901	420.058	38.6	12.59
ARIMA(1,1,2) (2,1,2)	.904	501.983	35.5	12.6
ARIMA(1,1,2)(2,1,3)	.916	483.016	34.6	12.47
ARIMA(0,1,2) (2,1,4)	.916	584.085	28.68	12.5

(Soruce: author won compiled for various model generated by Spss 17)

From the above models we further streamline the model by looking at the residual, in order to know that the model has white noise. If the model has white nose then we can used it for forecasting. The residual of ARIMA (0, 1, 1) (1, 1, 1) has white noise (see fig 5).

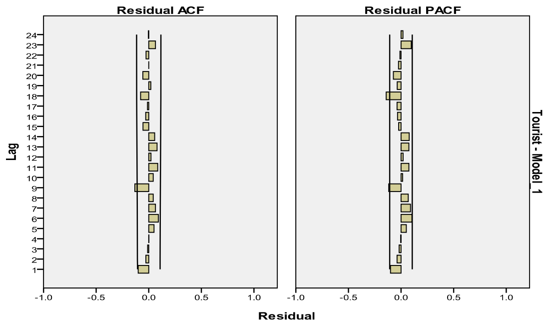

Figure 5: Source: Generated by spss17

3.4 Forecasting of Tourist for 2013 and 2014

Now, we turn to the application (i.e., forecasting). Our objective is to predict the 12 future values of time series (out-of-sample monthly forecasts for both time series). Table 3 shows monthly forecasted results with confidence limits for NA time series. As expected, March, April, and September October, November are the months with the most prominent values, thus expressing the extension of strong seasonal movement in the number of arrivals in Bhutan.

Table 2:

2013	*Tourist*	*2014*	*Tourist*
13-Jan	1358	14-Jan	1890
13-Feb	2274	14-Feb	2900
13-Mar	6865	14-Mar	8234
13-Apr	7752	14-Apr	9773
13-May	3909	14-May	4809
13-Jun	2218	14-Jun	3066
13-Jul	2170	14-Jul	2716
13-Aug	3416	14-Aug	4361
13-Sep	6772	14-Sep	8510
13-Oct	13291	14-Oct	16022
13-Nov	7846	14-Nov	9508
13-Dec	3241	14-Dec	4257
Total	61116	Total	76052

(Forecasted value from the model using Spss 17)

4. Conclusion and Recommendation

In this paper, we considered two forecasting models in order to determine the size of the flows of tourism demand in Bhutan according to the number of arrivals and overnight stays, which are some of the key variables in tourism analysis. Theoretical framework that we used was the Box-Jenkins methodology for seasonal ARIMA models. After constructing the appropriate models, we utilized them to generate the forecasts of demand within Bhutan tourism. The obtained results (i.e., forecasted values) can provide important information needed for an adequate destination. Other techniques like Winter Multiplicative may be effective as every October and November the tourist arrival is maximum.

References

[1] Akal, M. (2004). Forecasting Turkey's tourism revenues by ARMAX model. *Tourism Management*, Vol. 25, No.5, pp.565-580.

[2] Akal, M. (2010). Economic implications of international tourism on Turkish economy. *TOURISMOS*, Vol. 5, No.1, pp.131-152.

[3] Chu, F.L. (2008a). Forecasting tourism demand with ARMA-based methods. *Tourism Management*, Vol. 30, No.2, pp.740-751.

[4] Chu, F.L. (2008b). Analyzing and forecasting tourism demand with ARAR algorithm. *Tourism Management*, Vol. 29, No.6, pp.1185-1196.

[5] Coshall, J.T. (2009). Combining volatility and smoothing forecasts of UK demand for international tourism. *Tourism Management*, Vol. 30, No.4, pp.495-511.

[6] Frechtiing, D.C. (1996), *Practical Tourism Forecasting*, Butterworth-Heinemann, Oxford.

[7] Kon S.C(2002). Forecasting Tourist Demand to Singapore: A Modern Time Series Approach. School of Applied Economics Faculty of Business and Law, Victoria University

[8] Bigivic M.(2012), "Demand forecasting within Montegegrin Toursm using Box-Jenkins Methodology fir Seasonal ARIMA models", *Tourism and Hospitality Management*, Vol. 18, No.1 pp 1-18,

[9] Moriey, C.L. (1990), "What is Tourism? Definitions, Concepts and Characteristics, *"The Journal of Tourism Studies*, 1 (1), 3-8.

[10] Pankratz, A. (1983), *Forecasting with Univariate Box-Jenkins Models; Concepts and Cases*, John Wiley & Sons, New York.

[11] Wong, K., Song, H., Witt, S.F. & Wu, D.C. (2007). Tourism forecasting: to combine or not to combine? *Tourism Management*, Vol. 28, pp.1068-1078.

[12] http://www.unescap.org/ttdw/Publications/TPTS_pubs/pub_2478/pub_2478_fulltext.pdf, pp 111, date of access 3-03-2011

[13] http://www.bhutantimes.bt/index.php?option=com_content&task=view&id=1410&Itemid=79 access on 20-12-12

[14] http://www.wttc.org/eng/Tourism_News/Press_Releases/Press_Releases_2008/World_Citizens_%E2%80%93_a_global_partnership_for_Travel_and_Tourism/

Author Profile

Dr. Elangbam Haridev Singh received his B.sc. in Physics, M.BA. Degrees and PhD in Economics on the Topics 'Problems and Prospect of Marketing: A case of Handicraft Industries in Manipur' from Manipur University in the year 1998, 2000, and 2013 respectively. Currently he is working as a Senior Lecturer in Royal University of Bhutan since 2009.

Acknowledgment: 1. Dr. Pawan Kumar Sharma Associate Professor Delhi University, CPL GCBS, Royal University of Bhutan.