Intervention time-series model using transfer function forTourism arrival in Sri Lanka

B.Yogarajah
Department of Physical Science, Vavuniya Campus, University of Jaffna, Sri Lanka
yoganbala@vau.jfn.ac.lk

Introduction

Tourism plays a vital role in the economy of most of the countries in the world. Tourism is the main industry in Sri Lanka and its major source of foreign exchange and revenue. World Economic Forum (2017) reported that Spain, France and Germany were the top three most tourist-friendly destinations in the world. At the meantime, Sri Lanka has been ranked 64thin the Travel & Tourism Competitiveness Index (TTCI) out of 141 countries. While further improvements are needed, Sri Lanka is taking small but important steps in the right direction, the report said.

Sri Lankan tourism has had many drawbacks during the last three decades mainly due to the civil war in the country until 2009 afterwards the Sri Lankan tourism industry has been growing dramatically. At present, Sri Lanka has overcome all these drawbacks and the industry is leading to a new era. In such a situation the forecast of tourist's arrivals is important since it would be beneficial the tourism-related industries like airlines, hotels and other stakeholders to adequately prepare for the expected number of tourists in the coming year.

The data related to this collection from Annual Statistical Report of Sri Lanka Tourism-2016. The time span used for the study is from the year 1970 to 2017 with intervention after the year 2009.

Intervention analysis is the application of modelling procedures for incorporating the effects of exogenous forces or interventions in time series analysis. One special kind of ARIMA model with input series is called an intervention model or interrupted time series model. In an intervention of this model, the input series is an indicator variable that contains values one after the post-war of 2009 that occurrence of an event affecting the response series tourism arrival in Sri Lanka. This intervention model can be used both to model and forecast the response series.

The objective of the study is to introduce intervention time series analysis for fitting a most appropriate model to predict tourist arrivals in Sri Lanka by using intervention modelling method.

Literature review

In literature, many research studies can be found on forecasting tourist arrivals in national and international context references. Those studies have used different methodologies to forecast the tourist arrivals based on the nature of the data. Application of Time Series Intervention Modelling for Modelling and Forecasting, Cotton Yield(Mrinmoy Ray,2014), in this model the intervention being the introduction of Bt Cotton variety in the year 2002. A systematic approach to modelling and forecasting in the time domain are illustrated in the classic work of Box-Jenkins (1970) or later books such as Shummay (1989), Harvey (1994), Wei (1994), and Caynor (1991). Autoregressive integrated moving average (ARIMA) modelling is overall the most accurate method for forecasting international tourist arrivals (Cho, 2001; Chu, 1998; Loganathan&Yahaya, 2010; Chang et al., 2011; Dimitrios et al., 2012; Saayman et al., 2010). The study was done by P.C.Padhan has fitted Seasonal Auto-Regressive Integrated Moving Average (SARIMA) models for forecasting future tourist's arrivals in India. In this fitted models are compared mainly using Mean Absolute Percentage Error (MAPE) values. Another study on forecasting tourist arrivals to Australia has used Box-Jenkins' Auto-Regressive Integrated Moving Average (ARIMA) and SARIMA models in their study. The most popular modelling procedures are those where "intervention" detection and estimation is paramount. Box and Tiao (1975) pioneered this type of analysis in their quest to solve the Los Angeles pollution problem. According to the literature, empirical studies of time series behaviour of the post-war international tourist arrivals to Sri Lanka had been carried out using different modelling approaches. Such as the classical time series decomposition approach (Kurukulasooriya&Lelwala, 2014) with 96% forecasting accuracy, Box-Jenkin's modelling and Holt - Winter's Exponential Smoothing approaches (Gnanapragasam& Cooray, 2016(a) & 2016(b)) with nearly 90% and 88% forecasting accuracy respectively. In this new approach, the transfer function model has been developed for predicting international tourist arrivals to Sri Lanka with 91% forecasting accuracy.

Methodology

Intervention Model

There are several sources of information about interrupted time series designs that are intended for an audience of social scientists. The most comprehensive discussion of the interrupted time series design is provided by Glass, Willson, and Gottman in Design and Analysis of Time Series Experiments (1975). Box-Jenkins strategy for time series modelling with features for the identification, estimation and diagnostic checking, and forecasting steps of the Box-Jenkins method.

Suppose that the ARIMA model for X_t(the observed series) with no intervention isWith the usual assumptions about the error series $X_t - \mu = \frac{\theta(B)}{\varphi(B)} w_t$. Where $\theta(B)$ is the usual MA polynomial and $\varphi(B)$is the usual AR polynomial. Let Z_t = the amount of change at time t that is attributable to the intervention. By definition, Z_t = 0 before time T (time of the intervention). The value of Z_t may or may not be 0 after time T.

Then the overall model, including the intervention effect, may be written as

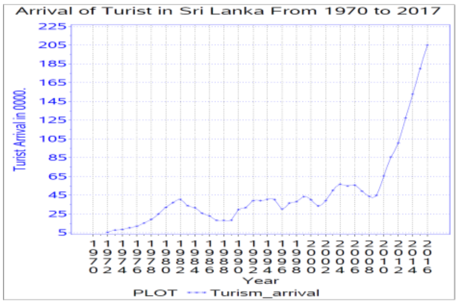
$$X_t - \mu = Z_t + \frac{\theta(B)}{\varphi(B)} w_t$$

Intervention analysis may be viewed as a type of regression analysis in which one or more predictor variables observed at equally spaced time points are postulated to have had an impact on a response variable observed at the same time points. In the simplest case, the intervention predictor variable is an indicator variable that takes the value zero before the intervention occurs and the value one after the intervention occurs. However, the predictor variables in an intervention analysis need not be indicator variables. We consider the general stationary intervention model (BoX and Tiao 1975) given by the intervention variables used are defined by X is equal one if year > 2009 otherwise zero.

Results and Discussions

In Figure 1(a), it is obviously observed that from the years 2009 there is a sudden upward trend look like in tourist arrivals. This is the beginning of the internal conflict in Sri Lanka. There are ups and downs in arrivals can be seen from the years 1983 to 2009. This is the period where the internal conflict took place in Sri Lanka. It reveals that there is a rapid increase in growth rate in total tourist arrivals after the post-war period. This is the reason for this study is to use Intervention time-series model using transfer function mainly focus on the international tourist arrivals to Sri Lanka.

By visual inspection of autocorrelation plot (ACF) and partial autocorrelation plot (PCAF) plots (Figure 1(b) and Figure 2(b)), the ACF decays very slowly shows how values of the series are correlated with past values of the series. The Augmented Dickey-Fuller Unit Root test (Table1) also results show that the tourist arrivals series is nonstationary since the series has a trend over time, seasonality, or some other nonstationary pattern. The usual solution of this result is to take the difference of the series from one period to the next and then analyze it. Tourist arrivals data consisted of heteroscedasticity errors. Therefore, Logarithm transforms and differenced the data were used to transform data to remove the heteroscedasticity errors.



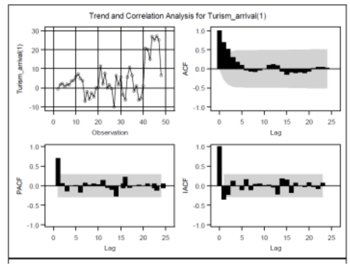
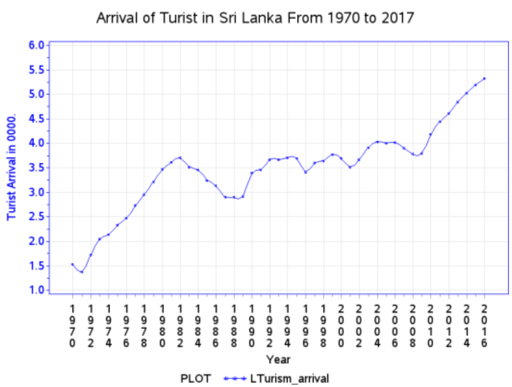


Figure 1. (a) Plot of yearly tourist arrivals

Figure 1(b). The plot of Difference, ACF and PACF



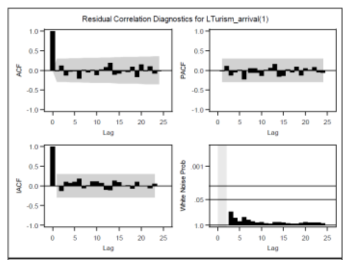


Figure 2(b): Plot of ACF and PACF

Figure 2(a): Plot of Log transformation of tourist arrivals

The Extended Sample Autocorrelation Function (ESACF) and Scanning methods(Table2) can tentatively identify the orders of autoregressive p=2,order of moving average q=1 and difference d=1 of ARIMA process with intervention based on iterated least squares estimates of the autoregressive parameters(Tsay and Tiao (1984)) and minimum values of Schwarzs Bayesian Criterion(BIC=3.9) Once an appropriate model has been identified and its parameters estimated, it is, of course, necessary to determine whether this choice is adequate. The diagnostic checks including the independence, normality and homoscedasticity of residuals are the most important stage of Intervention model building process. Since the visual inspection of *Figure 3*, the autocorrelation function plots indicates that these residual autocorrelation plots are white noise and normal probability plot shows a straight line, it is reasonable to assume that the observed sample comes from a normal distribution. This is the statistical evidence that the assumption that the random errors are an independent sample from a normal distribution. These diagnostics indicate that the model has produced only uncorrelated white noise, which is characteristic of a properly specified model.

Table1: Augmented Dickey-Fuller Unit Root Tests

Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr< Rho	Tau	Pr< Tau	F	Pr > F
Zero Mean	1	-16.0023	0.0034	-2.86	0.0051		
	12	-6.1878	0.0788	-1.06	0.2562		
Single Mean	1	-21.8734	0.0030	-3.25	0.0230	5.31	0.0380
	12	23.9424	0.9999	-1.75	0.3976	1.53	0.6867
Trend	1	-21.8682	0.0236	-3.21	0.0961	5.17	0.1838
	12	101.6854	0.9999	-2.27	0.4394	4.15	0.3748

Table 2: oder selection for the Interventional model

ARMA(p+d,q)Tentative Order SelectionTests					
SCAN			**ESACF**		
p+d	**q**	**BIC**	**p+d**	**q**	**BIC**
3	1	3.902	3	1	3.901709
1	2	3.960	0	2	4.851274
0	4	4.357	1	2	3.960112
			4	0	4.034265

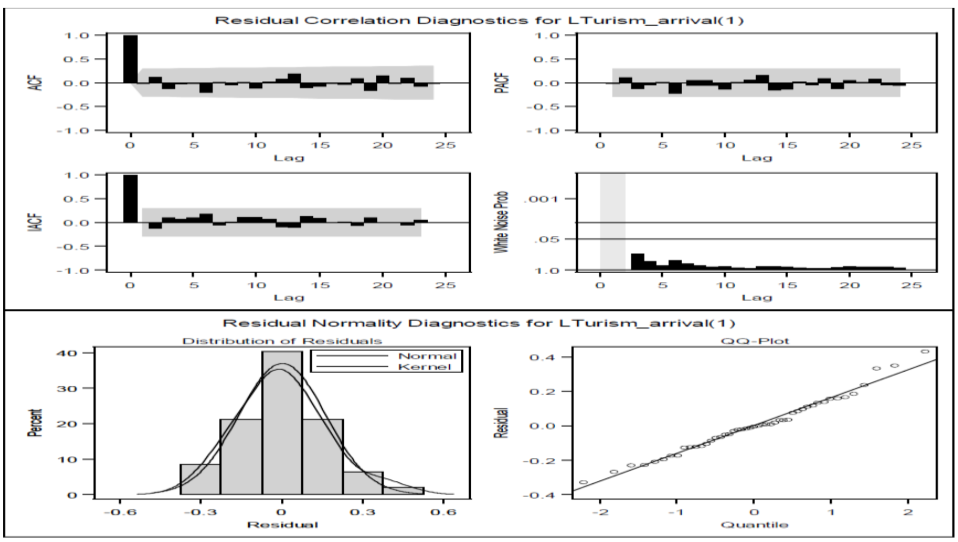


Figure 3: Plots of ACF,PACF, Normal Probability plot for Residual

Table 3 : Model Validation

ARIMA (p, d, q)	R-squared	RMSE	MAPE	AIC
ARIMA(2,1,1)	0.878	130.41	55.46	29.6
ARIMA(1,1,0)	0.578	125.70	58.29	45.2
ARIMA(1,1,1)	0.779	274.90	56.07	30.9
ARIMA(1,1,2)	0.703	719.18	48.45	31.2
ARIMA(2,1,2)	0.793	720.74	4261	31.7
ARIMA(3,1,1)	**0.623**	**115.35**	**38.35**	**32.4**

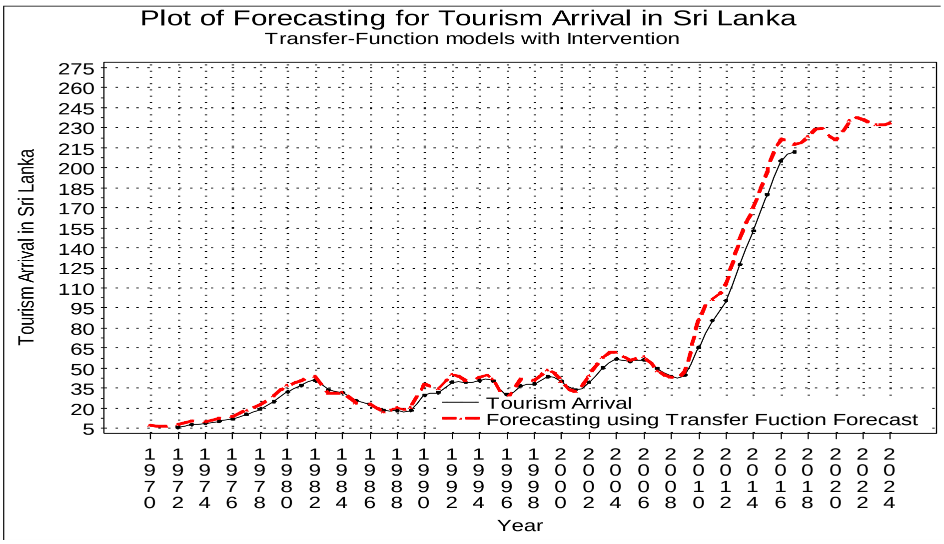


Figure 4: Plot of forecasting and real data for Tourist arrival to Sri Lanka

Conclusions and Recommendations

Key findings indicate that the tourist arrivals patterns in the study period are auto-regressive, as such they do depend on the past history of the data and combination of past random noises.

Analysis of performance of the tourist arrivals from preceding 47 years traded value gives us ARIMA (2, 1, 1) model with intervention post-war period which helps us in predicting the future values of tourist arrivals. ARIMA (2, 1, 1)with the intervention was chosen from six different model parameters by using minimum values of RMSE, MAPE and AIC(Table3), as it provides the best model which satisfies all the criteria of fit statistics while other five failed the fitted model is

$$(1-B)\log(Arrivals)_t = 0.067785 + 0.064917X + \frac{1\ +\ 0.17013\ B ** (1)\ -\ 0.28196\ B ** (2)}{1\ +\ 0.45655\ B ** (1)}\, a_t$$

Where B is the backshift operator and it is the random error.

The intervention time-series model using transfer function for Tourism arrival was identified as the appropriate model to Sri Lanka.

References

Harvey, A.C., & Phillips, G.D.A.(1979) *Maximum likelihood estimation of regression models with an Autoregressive moving average disturbances*Biometrika, 49-58.

The World Economic Forum Report -2017.

D. K. Ishara1* and P. Wijekoon1 2017, *Time Series Analysis for Modeling and Forecasting International Tourist Arrivals in Sri Lanka*.

Glass, Willson, and Gottman in *Design and Analysis of Time Series Experiments* (1975).

C. Lim and M. McAleer, *Time series forecasts of international tourism demand for Australia*, Discussion paper no 533, 2001.

P.C. Padhan 2010, *Forecasting tourists' footfalls in India: an assortment of Completing Models*, International journal of business and Management, Vol 6, No 5, 2010.

V. Cho 2003, *A comparison of three different approaches to tourist arrivals forecasting*, Tourism Management, Vol.24, No 3, 323-330, 2003. Box, G.E.P. and Tiao, G.C., 1975 pp. 70-79. *Intervention Analysis with Applications to Economic and Environmental Problems*. JASA.

Box Jenkins, *SARIMA application. South Asian Journal of Tourism and Heritage*. 2010; 3(2): 50-60

SAS/ETS(R) 9.2 User's Guide (2010), SAS Institute

Cantor, D. and K. C. Land (1985), *Unemployment and crime rates in the post-world War II United States: A theoretical and empirical analysis,* American Sociological Review 50:317-332.

Hultkrantz L, Olsson C: *Chernobyl effects on domestic and inbound tourism in Sweden - a time series analysis*. Environmental and Resource Economics. 1997, 9: 239-258.

Cho V 2001, *Tourism forecasting and its relationship with leading economic indicators*. Journal of Hospitality and Tourism Research. 2001; 24:399-420.

Loganathann&Yahayai. *Forecasting international tourism demand in Malaysia using Box Jenkins SARIMA application.* South Asian Journal of Tourism and Heritage. 2010; 3(2): 50-60.

Saayman A &Saayman M. 2010, *Forecasting tourist arrivals in South Africa*. ActaCommercii. 2010; 281-293.

Annual Statistical Report of Sri Lanka Tourism– 2016

Mrinmoy Ray, V. Ramasubramanian2014, *Application of Time Series Intervention Modelling for Modelling and Forecasting Cotton Yield*, Statistics and Applications Volume 12, Nos. 1&2, 2014 (New Series), pp. 61-70