Hybrid Trend – ARIMA Model for Forecasting Employment in Tourism Industry in Sri Lanka

Review of Integrative Business & Economics Research

K.M.U.B Konarasinghe
Institute of Mathematics & Management, Sri Lanka

ABSTRACT

Tourism is a fast-growing industry in Sri Lanka; as a result, the employment in the industry also shows a rapid growth. In light of that, the study was focused on identifying suitable forecasting techniques for total employment of tourism industry in Sri Lanka. Annual employment data for the period of 1970 to 2015 were obtained from the Sri Lanka Tourism Development Authority. The trend models; Linear, Quadratic, Growth Curve and Pearl Reed Logistic were tested. The Anderson-Darling test revealed the residuals of all four models were normally distributed, but Ljung-Box Q-test dies not confirm the independence of residuals. Therefore, the de-trended data were further analyzed; the stationary of the series was tested by Augmented Dickey-Fuller test. Then the Auto Regressive Integrated Moving Average (ARIMA) model was tested on each series. The ARIMA model was well fitted for de-trended series of Linear trend model and Growth Curve model. Hence, the residuals of two hybrid models; Linear trend-ARIMA and Growth Curve trend-ARIMA models were tested for model assumptions. It was concluded that both hybrid models; Linear trend-ARIMA and Growth Curve-ARIMA are suitable for forecasting total tourism employment.

Keywords: Trend Model, ARIMA, Hybrid Model

1. INTRODUCTION

The tourism industry is one of the largest and most dynamic industries in the global economy today. It generated 105,408,000 jobs directly in 2014 and it grow to 107,519,000 in 2015(WTTC, 2015).This includes employment by hotels, travel agents, airlines and other passenger transportation services. It also includes, for example, the activities of the restaurant and leisure industries directly supported by tourists. Tourism industry creates jobs, accounting 1 in 11 worldwide (WTTC, 2015). Direct and indirect employment opportunities are major segments of tourism. Direct employment includes, employed by hotels, travel agents, airlines and other passenger transportation services, agencies providing recreational facilities, tourist guides, tourist shops and other organizations in the state sector. Indirect employment is businesses which sell goods and services to the tourism sector.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

The Sri Lankan tourism market consist all regions in the world. Konarasinghe (2016a) showed that there is an increasing trend of tourist arrivals from all regions. It could be an evidence of a growth of tourism industry in Sri Lanka. The growth of industry ensures the growth of both direct and indirect employment opportunities in tourism. Figure 1 is the time series plot of the total growth of employment in the tourism industry in Sri Lanka.

Figure 1: Time Series plot of total employment

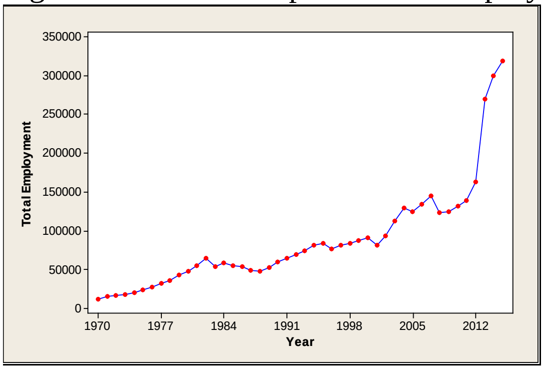


Figure 1 clearly shows the improvement in total employment over the years. The year 2009 is the significant landmark of Sri Lankan economy. After 2009 there is a growth of Sri Lankan tourism industry in many ways (SLTDA, 2013). As a result, total employment within the industry also shows a remarkable improvement.

1.1 PROBLEM STATEMENT

The growth of tourism industry and the growth of employment in the tourism industry are in line. Therefore, it is important to selecting and training individuals for specific job functions and charging them with the associated responsibilities. Konarasinghe (2015a) has shown the importance of planning and forecasting employment in the industry, in order to overcome the surplus and shortfall of employment. Forecasting helps for planning; planning helps to minimize the risk and maximize the benefits of the entire industry. It leads to the satisfaction of stakeholders of the industry. However, fewer attempts were found in forecasting employment in the tourism industry in Sri Lanka. Therefore, it is a timely requirement to the development of the industry in Sri Lanka.

1.2 OBJECTIVE OF THE STUDY

To identify the suitable time series techniques for forecasting total employment of tourism industry in Sri Lanka

1.3 SIGNIFICANCE OF THE STUDY

Researchers have paid attention in forecasting employment in many fields, but least attention is found in the field of tourism. The industry creates millions of job opportunities worldwide. Also, it provides solutions to unemployment in Sri Lankan context. As such, it is important to forecast the employment of tourism industry. The results of the study will be very helpful to estimate the future numbers of employees required and the likely skills and competencies they will need. It will be useful to achieve business objectives, identifying the types and quantities of skills and developing

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

proactive strategies. It will be useful for workforce planning in both public and the private sector in the tourism industry. Further, it is useful for developing various training programs such as workshops, academic and professional courses related to hospitality management.

2. LITERATURE REVIEW

The literature review of the study was focused on forecasting employment in various tourism destinations. However, there were fewer attempts in that light, in Sri Lankan context as well as the other destinations. Therefore, the author of the study reviewed the studies related to employment forecasting on various industries across the world.

Chau (1970) construct an economic model, for making short-run forecasts on civilian personal income and employment of Hawaii. He used multiple regression models for forecasting. He concluded that the forecasting ability of the model is quite accurate in employment forecasting.
Witt, Song, and Anhill (2004) forecasted tourism generated employment in Denmark. They used Autoregressive Distributed Lag Models (ADLM) approach. The ADLM were well fitted for the purpose and measurements of errors were satisfactorily small. Rapach and Strauss (2005) forecasted employment growth in Missouri using the ADLM approach. The performance of their model also was highly satisfactory.
Paquet, Sargent, and James (2006) examine the past and future behavior of the employment rate of men and women in Canada using log –linear model. Model validation and verification confirmed that the projections of fitted models are reliable for short-term forecasting, but the same was not true for in the long run.
Patuelli, Reggiani, Nijkamp and Blien (2006) forecast regional employment in both the former West and East Germany by using the neural network. The results revealed that the absolute measurements are not up to the level of satisfactory, but in contrast, the relative measurements are highly satisfactory. Patuelli, Reggiani, Nijkamp and Schanne (2011) also developed neural network models for regional employment forecasting in three districts in German. They concluded that the performances ability of the neural network is highly satisfactory.
Vitartas, and Ford(2008) forecast employment demand in local government area in Northern New South Wales, Australia by using Logistic Regression. The forecasting ability of the model was low in their study.
Chang and Sung (2010) forecast industry employment for a resource-based economy in a state of Georgia. They developed Bayesian vector autoregressive models for the purpose. They concluded that the fitted models perform well in long run. Raoufinia (2016) used Vector Auto-Regressive models to forecast employment growth in Sweden. The results of the study were satisfactory up to a certain extent in short-term forecasting. Konarasinghe (2015b) tested linear and non – linear trend models for forecasting direct employment trend in the tourism industry in Sri Lanka. The results of this study confirmed that linear trend model is suitable for forecasting.

According to the literature; researchers have tested statistical and soft computing techniques as forecasting tools. Most of them have used multivariate techniques, while some of them used univariate techniques. The ADLM and Regression models were the commonly tested models. The performances of the models were satisfied. It was hard to find any hybrid forecasting techniques in forecasting employment.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

3. METHODOLOGY

An annual employment data for the period of 1970 to 2015 were obtained from annual reports of Sri Lanka Tourism Development Authority (SLTDA). Box and whisker plot was used for outlier detection in the data set. Figure 1 is the time series plot of total tourism employment. It is clear that there is an increasing trend. Therefore the trend models; Linear, Quadratic, Growth Curve and Pearl Reed logistic were tested on log transformed series. Four trend models were tested with log transformation. They are;

Linear trend model: $\ln Y_t = \alpha + \beta \text{ t} + \varepsilon$ (1)

Quadratic trend model: $\ln Y_t = \alpha + \beta_1 \text{t} + \beta_2 \text{t}^2 + \varepsilon$ (2)

Growth curve model: $\ln Y_t = \alpha(\beta^t) + \varepsilon$ (3)

Pearl – Reed Logistic model: $\ln Y_t = \dfrac{\text{A}}{\alpha + \beta(\gamma^t)} + \varepsilon$ (4)

Trend analysis revealed that the residuals of the models were not independent. Therefore de-trended data were analyzed and it was intended to model them. The Auto Regressive Integrated Moving Average (ARIMA) model was tested for the purpose.

An ARIMA model is given by:

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \quad (5)$$

Where; $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 .... \phi_p B^p$

$\theta(B) - 1 - \theta_1 B - \theta_2 B^2 ...... \theta_P B^q$

$\varepsilon_t =$ Error term

D = Differencing term

B = Backshift operator ($B^a Y_t = Y_{t-a}$)

The time series plots and Auto Correlation Function (ACF) used for pattern recognition. The Anderson-Darling test, Ljung-Box Q (LBQ) test, ACF and Augmented Dickey-Fuller (ADF) test were used to test normality, independence of residuals and stationary of the series. Forecasting ability of the models was assessed by considering Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE) and Mean Absolute Deviation (MAD).

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

4. RESULTS

Data analysis is organized as follows;

4.1. Trend Analysis

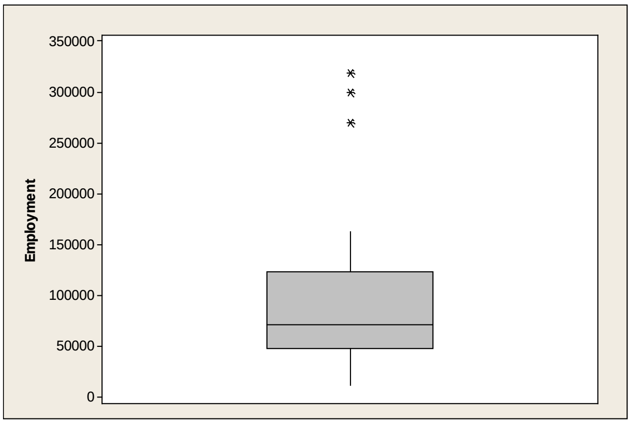
4.2. Model De-trended series

4.3. Testing the Validity of Hybrid models

Outliers are the extremely large or small values of a data set. They were identified with the help of Box Plot (Figure 2) and replaced by moving an average of order three. The study adopted the technique used by Konarasinghe, Abeynayake, and Gunaratne (2016) and Konarasinghe (2016b) for outlier adjustment. That is if the i^{th} value of a series is an outlier;

$$i^{th} value = [(i-1)^{th} value + (i-2)^{th} value + (i-3)^{th} value]/3$$

Figure 2: Box plot of Total Employment



4.1 TREND ANALYSIS

The trend models; Linear, Quadratic, Growth Curve and Pearl Read Logistic were tested. The relative and absolute measurements of errors of all fitted models are small. The Anderson-Darling test confirmed the normality of residuals, but the LBQ-test did not confirm the independence of residuals. Therefore none of the trend models can be recommended for forecasting.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Table 1: Model summary of trend models

Model	Model Fitting	
$\ln Y_t = 9.89346 + 0.0512439t$	MAPE	1.50424
	MAD	0.15990
	MSE	0.04509
	Normality	P= 0.596
	Independence of Errors	No
$\ln Y_t = 9.57093 + 0.0962478t - 0.00107152\,t^2$	MAPE	1.25241
	MAD	0.13586
	MSE	0.02712
	Normality	P = 0.104
	Independence of Errors	No
$\ln Y_t = 9.91080\ (1.00476^t)$	MAPE	1.57335
	MAD	0.16756
	MSE	0.04952
	Normality	P= 0.447
	Independence of Errors	No
$\ln Y_t = \frac{10^2}{7.12261 + 2.91776(0.980952^t)}$	MAPE	1.43602
	MAD	0.15195
	MSE	0.04229
	Normality	P =0.081
	Independence of Errors	No

4.2 MODEL DE-TRENDED SERIES

The de-trended data of each trend model was extracted and plotted. Figure 3 is the time series plot of de-trended data, extracted from linear trend model. Figure 4, Figure 5 and Figure 6 are the time series plots of de-trended series of Quadratic trend model, Growth Curve trend model and Pearl Reed Logistic trend model respectively. As they showed wave-like patterns, they were tested for stationary by ADF test. The ADF test confirmed the stationary of the series or differenced series. Then the ARIMA (p, d, q) model was tested on them. ARIMA (1, 0, 0) was the best fitting model for de-trended series of; Linear, Growth Curve and Pearl Reed Logistic trend models.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Figure 3: De-trended series of Linear trend trend model

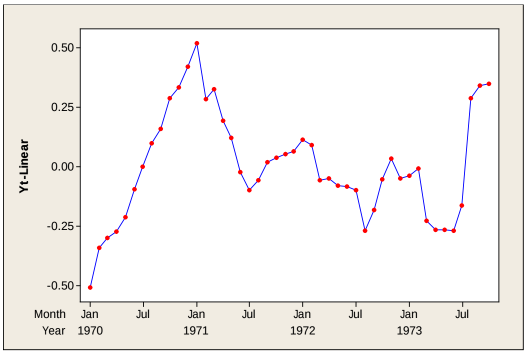


Figure 4: De-trended series of Quadratic model

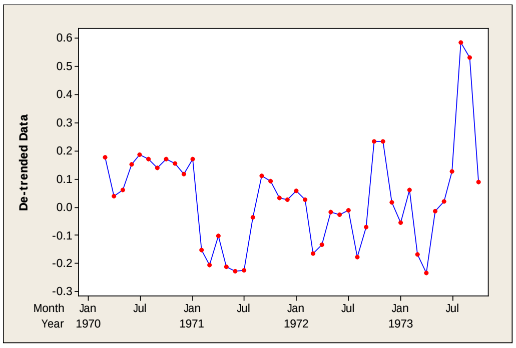


Figure 5: De-trended series of Growth Curve trend model

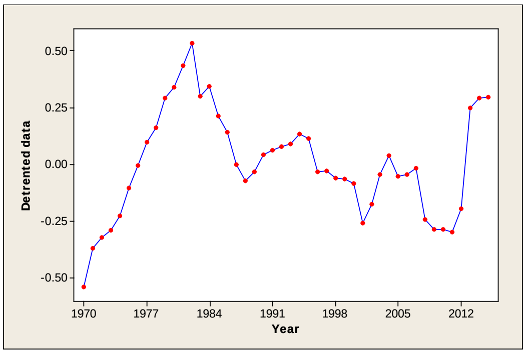
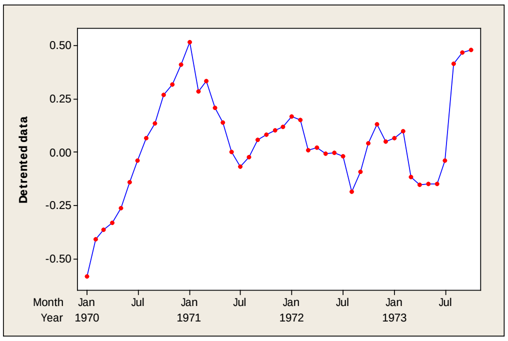


Figure 6: De-trended series of Pearl Reed trend model



The ARIMA (0, 2, 1) was the best fitting model for de-trended series of Quadratic trend model. Table 2 is the summary of best-fitting ARIMA models. The residuals of ARIMA (1,0,0); fitted to de-trended data of Linear and Growth Curve series were normality distributed and independent, but the fitted models of de-trended series; Quadratic and Pearl Reed Logistic models were not normally distributed.

Table 2: Summary of ARIMA models

De-trended data of;	Best Fitting Model	Model Fitting		Model Verification		Remarks of Residuals
		MSE	MAD	MSE	MAD	
Linear model	ARIMA (1,0,0)	0.0088	0.0751	0.2570	0.3703	Normal Independent
Quadratic model	ARIMA (0,2,1)	Not Fitted				Not Normal, Independent
Growth-Curve	ARIMA (1,0,0)	0.0091	0.0762	0.2388	0.3570	Normal, Independent
Pearl Reed	ARIMA (1,0,0)	Not Fitted				Not Normal, Independent

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Therefore, ARIMA model can be used to model the de-trended series of Linear and Growth Curve models. Measurements of errors of these two models were also satisfactory small.

4.3 TESTING THE VALIDITY OF HYBRID MODELS

Based on the results of 4.2 and 4.3 of the analysis, following two hybrid models were selected for forecasting;

Linear- ARIMA model: $Y_t = \alpha_1 + \beta t + \alpha_2 + \sum_{i=1}^{p} \varphi_i Y_{t-i} + \varepsilon_t$ (6)

Growth Curve -ARIMA model: $Y_t = \alpha_1 + \beta^t + \alpha_2 + \sum_{i=1}^{p} \varphi_i Y_{t-i} + \varepsilon_t$ (7)

Forecasts were estimated for both hybrid models; hence the residuals were calculated. The residuals of both models were normally distributed and independent. Measurements of errors of both models were very small.

Table 3: Summary of hybrid models

Model	Model Fitting		Model Verification	
$\ln Y_t = 9.89 + 0.05t + 0.96 \ln Y_{t-1}$	MAPE	0.6894	MAPE	3.0703
	MAD	0.0751	MAD	0.3857
	MSE	0.0088	MSE	0.2768
	Normality	P = 0.062		
	Independent	Yes		
$\ln Y_t = 9.90(1.00^t) + 0.978 Y_{t-1}$	MAPE	0.7002	MAPE	2.7834
	MAD	0.0762	MAD	0.3496
	MSE	0.0091	MSE	0.2278
	Normality	P = 0.075		
	Independent	Yes		

Figure 7 and Figure 8 are the time series plot of actual vs fits of above two models. These figures clearly show that the actual data and fits are very close. Hence, these two models are selected as suitable models for forecasting total tourism employment of Sri Lanka;

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Figure7: Actual Vs Fits Linear-ARIMA

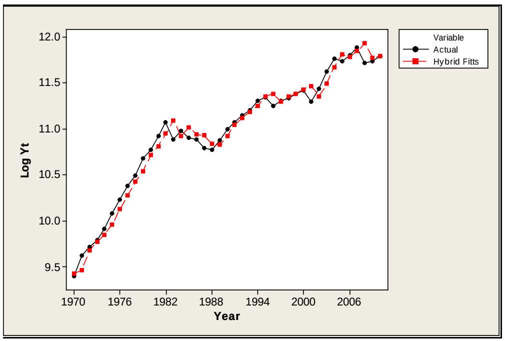
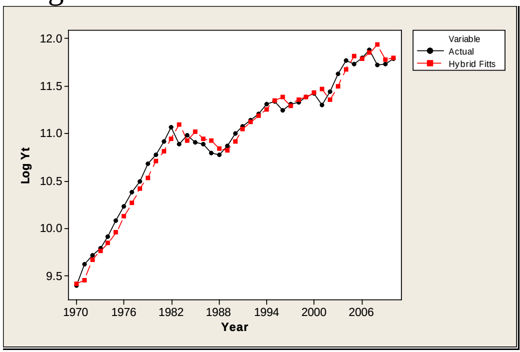


Figure 8: Actual Vs Fits GC-ARIMA



5. CONCLUSION AND RECOMMENDATION

Tourism creates millions of jobs, all over the world. The tourism industry is identified as a solution provider for unemployment. It is true for the Sri Lankan context as well. However, there exists a knowledge gap in forecasting tourism employment. The study was focused on finding the suitable techniques for forecasting total employment of the tourism industry in Sri Lankan context. Firstly, the trend models were tested for the purpose, and then the modeling was extended for the de-trended series. The results of the study revealed that the hybrid models; Linear trend - ARIMA and Growth Curve trend –ARIMA models are suitable for forecasting total tourism employments of the country.

It is clear that the employment in the industry is increasing, but it is not clear the types of jobs are in high demand. If the forecast can be made on job category wise, then it would help identifying job specific requirements. It would be a definite advantage for planning and policy making. Therefore, it is recommended to extend the study on job specific information; both direct and indirect employment in the tourism industry of Sri Lanka.

REFERENCES

[1] Chau, C.L. (1970) "An econometric model for forecasting income and employment in Hawaii", *Economic Research Center*, University of Hawaii.

[2] Chang, K. S., & Sung K. A. (2010) "Forecasting industry employment for a resource-based economy using Bayesian Vector Autoregressive Models", *The Review of Regional Studies*, 40(2), 181–196.

[3] Konarasinghe, K.M.U.B. (2015) "Trend Analysis of direct employment in the tourism industry of Sri Lanka", *Conference Proceedings of the 4th International Conference of the Sri Lankan Forum of University Economists*, 31.

[4] Konarasinghe, K.M.U.B. (2016) "Time Series Patterns of Tourist Arrivals to Sri Lanka", *Review of Integrative Business and Economics Research*, 5(3), 161-172.

[5] Konarasinghe, W.G.S., Abeynayake, N.R., & Gunaratne, L.H.P. (2016) "Circular Model on Forecasting Returns of Sri Lankan Share Market", *International Journal of Novel Research in Physics, Chemistry, and Mathematics*, 3(1), 49-56.

[6] Paquet, M.F., Sargent, T.C, & James, S. (2006) "Forecasting employment rates: a Cohort approach". Department of Finance, Working Paper.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

[7] Patuelli, R., Reggiani, A., Nijkamp, P., & Blien, U. (2006) "New neural network methods for forecasting regional employment: an analysis of German labor markets", *Spatial Economic Analysis,* 1 (1), 7–30.

[8] Patuelli, R ., Reggiani, A, Nijkamp, P., & Schanne, N. (2011) "Neural networks for regional employment forecasts: are the parameters relevant?", *Journal of Geographical Systems*, 13 (1), 67–85.

[9] Raoufinia, K. (2016) "Forecasting employment growth in Sweden using a Bayesian VAR model", *National Institute of Economic Research*, Working paper (144).

[10] Rapach, D. E., & Strauss, J, K. (2005) "Forecasting employment growth in Missouri with many potentially relevant predictors: an analysis of forecast combining methods", *Federal Reserve Bank of St. Louis Regional Economic Development*, 1(1), 97-112.

[11] Sri Lanka Tourism Development Authority (SLTDA). (2013) "Summary report 2013", Online at http://www.sltda.lk/statistics.

[12] Vitartas, P., & Ford, L., (2008) "Forecasting employment demand: a comparison of employers' perceptions and historical data in a regional LGA", *Journal of Economic and Social Policy,* 12(2).

[13] Witt, S.F., Song, H, & Anhill, S.W. (2004) "Forecasting tourism-generated employment: the case of Denmark", *Tourism Economics*, 10 (2), 167–176.

[14] World Travel and Tourism Council (2015). "Travel & tourism economic impact 2015 world", Online at http://www.wttc.org/.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)