International Journal of Research and Review
www.gkpublication.in E-ISSN: 2349-9788; P-ISSN: 2454-2237

Original Research Article

Patterns of Tourist Arrivals to Sri Lanka from Asian Countries

Konarasinghe Mudiyanselage Udaya Banda Konarasinghe

Institute of Mathematics and Management, Nugegoda. Sri Lanka.

Received: 15/11/2016 *Revised: 18/11/2016* *Accepted: 18/11/2016*

ABSTRACT

The pattern recognition is the key to success in time series forecasting. This study was focused on pattern recognition of tourist arrivals from leading Asian countries to Sri Lankan tourism market. Monthly time series data from January 2008 to December 2014 were used in this study. The regions selected for the study were the top five in market position in the Asian region. They are; India, Maldives, China, Japan, and Pakistan. Descriptive statistics, Time Series plots and Auto-Correlation Functions (ACF) were used for pattern identification and one way- Analysis of Variance (ANOVA) was used for mean comparison of tourist arrivals from selected countries. The average arrivals from India and Maldives were 13036 and 4000 consecutively. A number of arrivals from India were normally distributed. Data series of all five were non-stationary. There is a significant difference of tourist arrival from India and Maldives compared to other Asian countries. The arrivals of India, Maldives, Japan and Pakistan show the seasonal patterns. It is recommended to test Moving Average methods, Exponential Smoothing techniques, Holt's Winter's three parameter models, Decomposition techniques, Seasonal Auto Regressive Integrated Moving Average (SARIMA), Circular Model, linear and non-linear trend models for forecasting arrivals in short term and long term.

Keywords: Auto-Correlation Function; Analysis of Variance; Non-stationary.

INTRODUCTION

Sri Lanka claims to a long history of over three thousand years, with one of the longest documented histories in the world. Sri Lanka consists of rich culture with many different ethnic communities, cultural, linguistic and religious diversities. Sri Lanka is loaded with lush tropical forests, white sandy beaches and panoramic and unique landscapes with rich biodiversity. It could be the attraction for tourism to Sri Lanka. [1,2] It is considered as a wonderful tourist destination in the world known by Asian poets, nothing the geographical location of the island and lauding its beauty, is called the "pearl upon the brow of India".

Sri Lanka entered the international tourism market in the late 1960s, with the introduction of Ceylon Tourism Development Act of 1966, Ceylon Hotels Corporation Act of 1968 together with the introduction of Ceylon Tourist Board and the Ceylon Tourism Plan in 1967. [3-1] Sri Lankan tourism industry has boomed to a new milestone of 1,527,153 arrivals in 2014.It increase of 19.8% compared to arrivals 1,274,593 of 2013. [3-2]

The peaceful environment of the country after the year 2009 could be the reason of increasing tourist arrivals to Sri Lanka. The tourism market of Sri Lanka consist all regions in the world. The Asian region is the top tourist producer to Sri Lankan tourism market every year. India, Maldives, China, Japan, and Pakistan are the top Asian countries produced tourist to Sri Lanka in 2014.

Problem statement

Tourist arrivals from Asian countries may have similar or different patterns. This

is a very complicated situation in forecasting arrivals. Therefore identifying the similarities and differences of the patterns of tourist arrivals in various countries in Asian region are important. It facilitates to spotting suitable time series techniques for forecasting. This study was concerned with pattern recognition of tourist arrivals from various countries from Asian region.

Significance of the study

Sri Lankan tourism industry is the third highest foreign income generator [3] Proper planning and controlling are essential to gain maximum benefits and minimize the risk of any industry. Forecasting plays a vital role in planning. The role of statistical modeling facilitates in forecasting of a large number of fields. [4] The pattern recognition of arrivals provides guidelines to spot the suitable univariate forecasting techniques for tourist arrivals to Sri Lanka from leading Asian countries. Findings of this study will be a lighthouse to develop new forecasting techniques. And also, it will help to spot the suitable techniques for forecasting arrivals. It will ensure the proper supply and demand management to minimize the wastage strategic resources.

Objectives of the study

- To identify the patterns of arrivals to Sri Lanka, from Asian countries
- To compare the number of arrivals to Sri Lanka from Asian countries

LITERATURE REVIEW

Studies based on pattern recognition

Pattern recognition of arrivals was an immense interest of the researchers. An algorithm used on the dynamic time warping technique for pattern recognition for stock prices or NASA telemetry data. [5] Pattern Modeling and Recognition System (PMRS) as a tool for pattern recognition in the financial industry. The study utilized time series data of six different financial indices. They are German DAX, British FTSE, French FRCAC, SWISS, Dutch EOE and the US S&P series. [6]

Detecting of any recurrent pattern in ECG is another vital area in medicine. Neural Networks and wavelet transformed templates used for pattern recognition. [7] Another study used Pattern recognition algorithms, time series representations, and dimensionality reduction techniques to find the behavior of the occurrence of medium and large earthquakes. [8] This study utilized tectonic, geological, seismic and gravimetric data from 27 areas of Spain and Portugal. Pattern recognition algorithms, time series representations, and dimensionality reduction techniques were the techniques used by them. Agglomerative hierarchical clustering for recurring pattern recognition used some of the researchers. [9] There is a review focuses on several applications for pattern recognition; process, design cycle and models. They emphasize that pattern recognition systems can be designed using template matching, statistical methods, syntactic methods and neural network. And also, they introduce the fundamentals of statistical pattern recognition from several applications. They conclude that it is an interdisciplinary subject, covering in the areas of statistics, engineering, artificial intelligence, computer science, psychology, and physiology, among others. [10] Another study emphasized the importance of Symbolic Aggregate approximation (SAX) in pattern recognition of a time series. The recognition of recurring patterns within multivariate time series was the purpose of the study of Spiegel. [11] Box- plots, Time series plots and Auto-Correlation Functions (ACF) are some of the pattern recognition tools used by some researchers. [12] The ACF's and time series plots used as pattern recognition tools in tourist arrivals to Sri Lanka from all regions. [13-1] Though tourism is a growing industry, there were least attempts made for studies based on pattern recognition of tourism growth in Sri Lanka and other countries as well. There is only one study was found related to pattern recognitions of tourist arrivals to Sri Lanka

from all regions. But it is hard to find a study focused on pattern recognition of each country within the regions.

MATERIALS AND METHODS

Norbert Wiener defines, "A pattern is essentially an arrangement. It is characterized by the order of the elements of which it is made, rather than by the intrinsic nature of these elements." [14-1] A pattern recognition process has been gaining benefits from long term and it facilitates to developed highly sophisticated skills and techniques to take actions according to what they observed at present under any situation. [14-2] It provides direction to solve problems in many fields; medicine, finance, engineering agriculture and much more. Tourist arrivals from Asian region were the population of the study. A sample of five countries with the highest number of arrivals was selected for the analysis. They were; India, Maldives, China, Japan, and Pakistan. Monthly tourist arrival data from January 2008 to December 2014 was obtained from annual statistical reports from 2008 to 2014, published by Sri Lanka Tourism Development Authority (SLTDA). Descriptive statistics used for representing the summary of results. Time series plots and ACF were used for pattern identification and one way- Analysis of Variance (ANOVA) was used for mean comparison of tourist arrivals from selected countries.

Statistical Methods

Auto Correlation Function (ACF)

Autocorrelation computes and plots the autocorrelations of a time series. Autocorrelation is the correlation between observations of a time series separated by k time units. The plot of autocorrelations is called the Auto Correlation Function or ACF.

Analysis of variance (ANOVA)

Analysis of variance (ANOVA) is similar to regression in that it is used to investigate and model the relationship between a response variable and one or more independent variables. However, analysis of variance differs from regression in two ways: the independent variables are qualitative (categorical), and no assumption is made about the nature of the relationship (that is, the model does not include coefficients for variables). In effect, analysis of variance extends the two-sample t-test for testing the equality of two population means to a more general null hypothesis of comparing the equality of more than two means, versus them not all being equal.

Normal plot of residuals

The points in this plot should generally form a straight line if the residuals are normally distributed. If the points on the plot depart from a straight line, the normality assumption may be invalid. If the data have fewer than 50 observations, the plot may display curvature in the tails even if the residuals are normally distributed. As the number of observations decreases, the probability plot may show substantial variation and nonlinearity even if the residuals are normally distributed, it is necessary to use the Anderson-Darling statistic, to assess whether the residuals are normally distributed. If the p-value is lower than the chosen a-level, the data do not follow a normal distribution.

RESULTS

Data analysis is organized region wise as follows;

1. Descriptive statistics of tourist arrivals.
2. Pattern recognition of tourist arrivals.
3. Comparison of tourist arrivals.

Outliers are the extremely large or small values of a data set. They were identified with the help of Box Plot (Figure 1) and replaced by moving an average of order three. That is, if the i^{th} value of a series is an outlier;

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

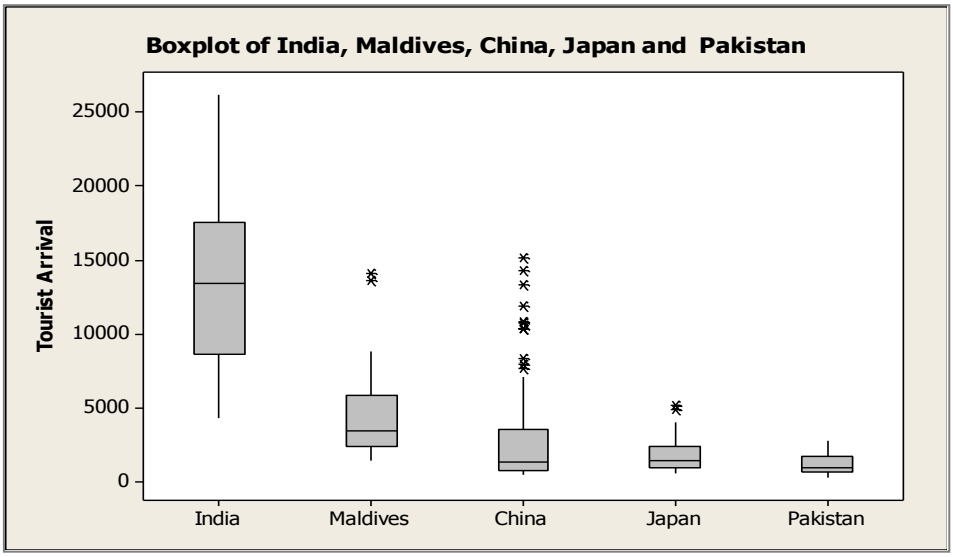

Figure 1: Box plot of arrivals from all countries

Descriptive statistics of tourist arrivals

Descriptive statistics for a number of arrivals were obtained country wise; it includes graphical summary along with measures of location, dispersion and the normality of the series.

Tourist arrivals from India

Graphical summary of descriptive statistics is shown in Figure 2. Minimum arrivals recorded from India were 4301 whereas maximum were 26153 during the period. The first quartile of arrivals is 8648. It means ¼ of the months had at most 8648 arrivals.

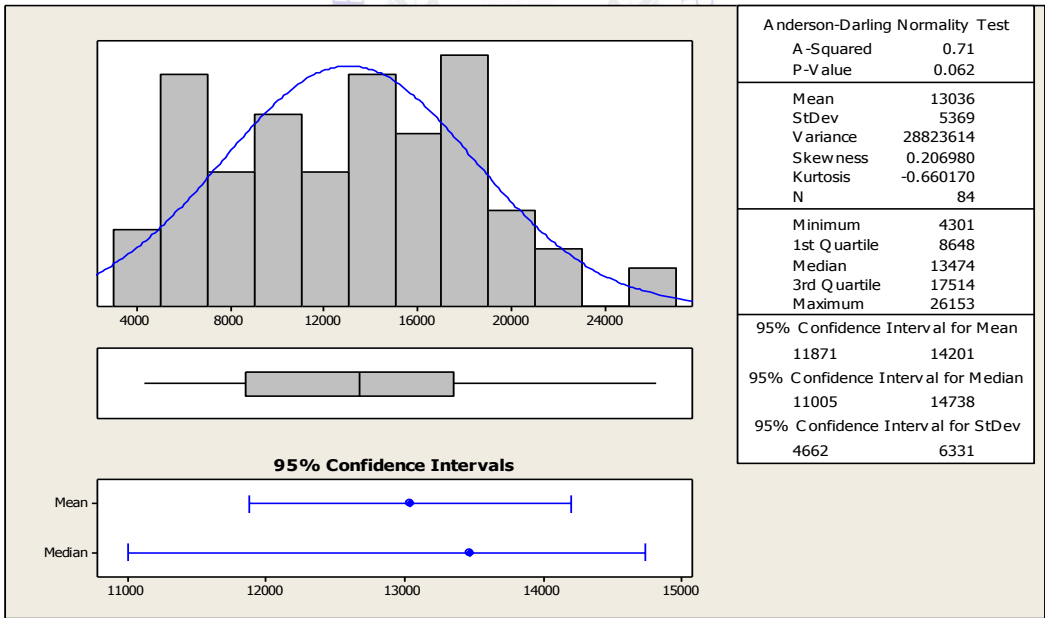

Figure 2: Graphical summary of tourist arrivals from India

A median arrival is 13474 and the third quartile of arrival is 17514. Histogram of the arrivals looks somewhat symmetrical. P value of the Anderson-Darling test (P= 0.062) is greater than the significance level (0.05). As such number of arrivals follows the normal distribution.

Tourist arrivals from Maldives

The Figure 3 shows that maximum tourist arrivals from Maldives were 8800 and minimum were 1452. The first quartile

of arrivals is 2441.A median arrival is 3503 and the third quartile of arrival is 5835.Histogram of the arrivals does not look symmetrical. P value of the Anderson-Darling test is less than the significance level (p-value <0.005). As such number of arrivals does not follow the normal distribution. They are positively skewed.

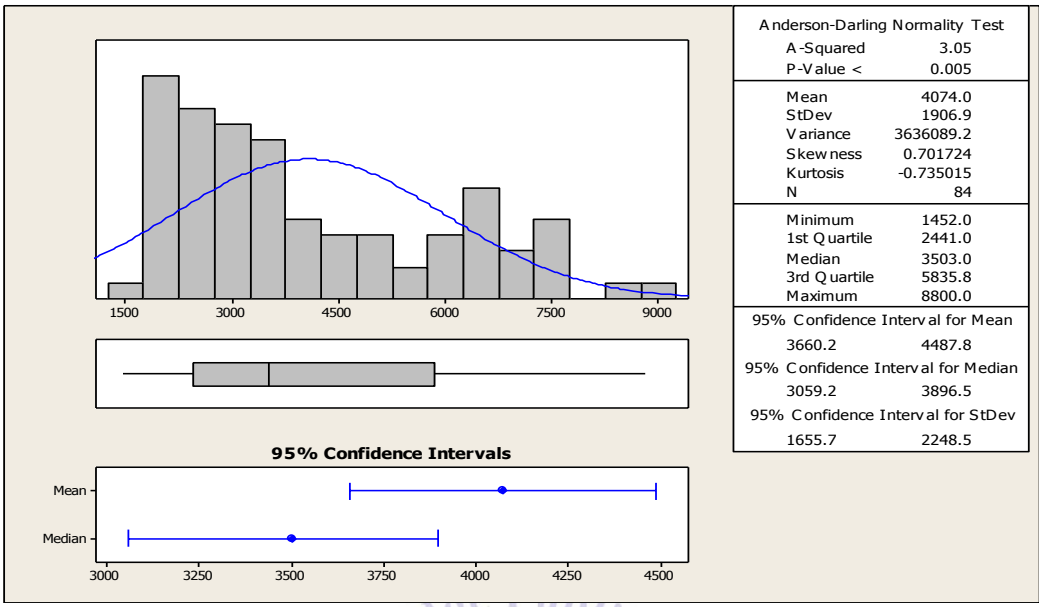

Figure 3: Graphical summary of tourist arrivals from Maldives

Tourist arrivals from China

Graphical summary of descriptive statistics is shown in Figure 4. Minimum arrivals recorded from China were 522 where maximum 7595.The first quartile of arrivals is 774. It means ¼ of the months had at most 774 arrivals. A median arrival is 1347 and the third quartile of arrival is 3542.Histogram of the arrivals does not look symmetrical. P value of the Anderson-Darling test is less than the significance level (p-value <0.005). As such number of arrivals does not follow the normal distribution. They are positively skewed.

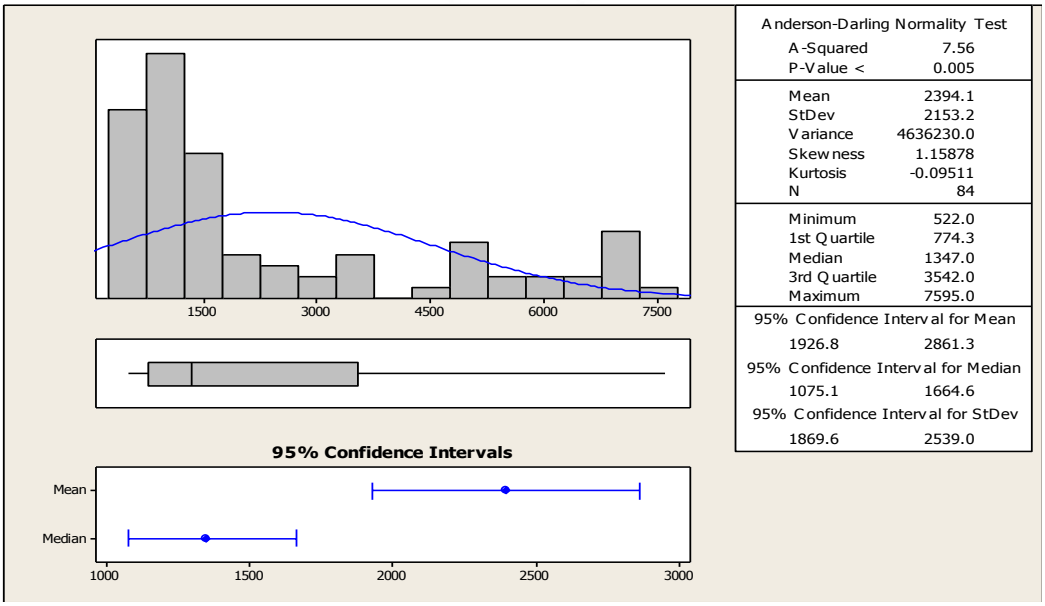

Figure 4: Graphical Summary of Tourist arrivals from China

Tourist arrivals from Japan

The Japan is the fourth of leading tourist producer from the Asian region.

According to Figure 5 of the graphical summary of descriptive statistics, minimum tourist arrivals recorded from Japan were 597 where maximum 4041. The first quartile of arrivals is 963. A median arrival is 1477 and the third quartile of arrival is 2380.Histogram of the arrivals does not look symmetrical. P value of the Anderson-Darling test is less than the significance level (p-value <0.005). As such number of arrivals does not follow the normal distribution. They are positively skewed.

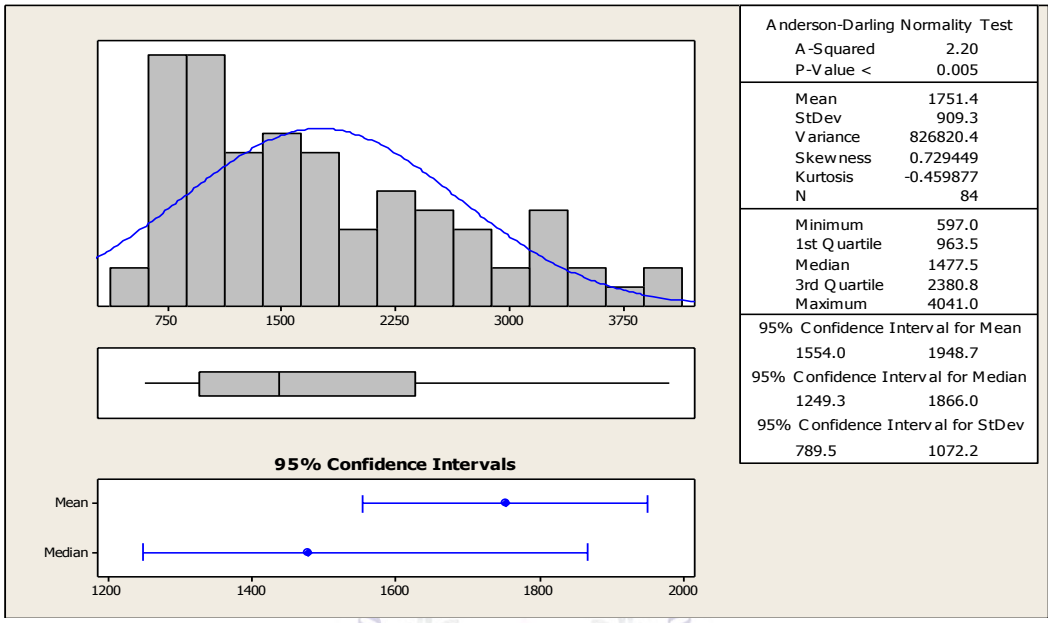

Figure 5: Graphical summary of tourist arrivals from Japan

Tourist arrivals from Pakistan

The Pakistan is the fifth leading tourist producer from the Asian region. According to Figure 6 of the graphical summary of descriptive statistics, minimum tourist arrivals recorded from Pakistan were 322 where maximum 2813. The first quartile of arrivals is 672. A median arrival is 993 and the third quartile of arrival is 1770. Histogram of the arrivals does not look symmetrical. P value of the Anderson-Darling test is less than the significance level (p-value <0.005). As such number of arrivals does not follow the normal distribution. They are positively skewed.

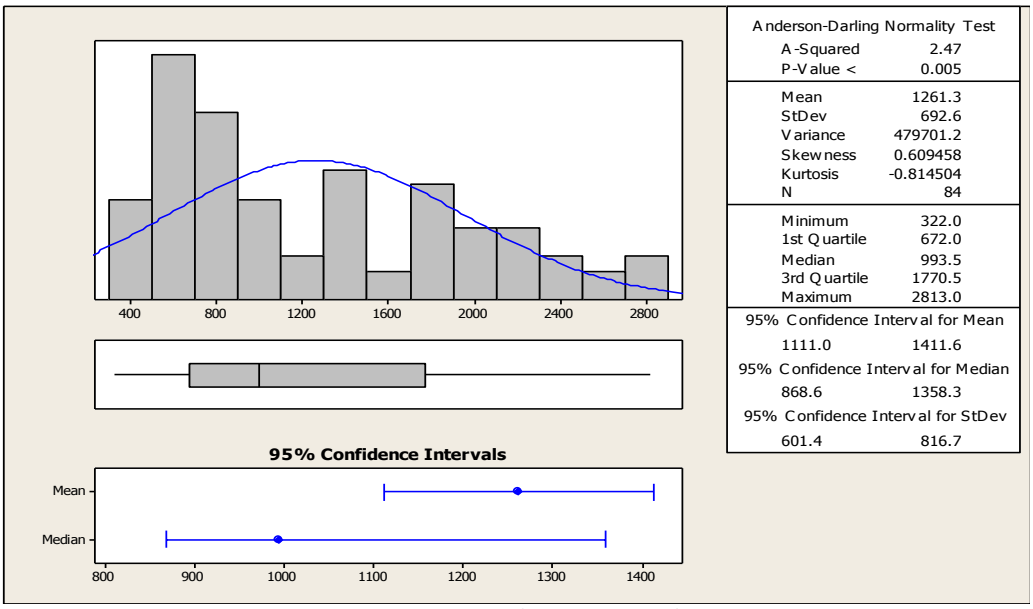

Figure 6: Graphical summary of tourist arrivals from Pakistan

Pattern recognition of tourist arrivals

Time series plots and ACF's were obtained, country wise;

Tourist arrival pattern of India

Time series plot of arrivals (Figure 7), shows the tourist arrival pattern of India. The behavior of the series clearly shows that there is an increasing trend from India. A series which has not constant mean and variance is known as non- stationary series. The ACF can be used to test the stationary of a series. Figure 8 is the ACF of arrivals from India confirmed that the series is not stationary. The pattern of ACF suggests seasonal behavior in arrivals from India.

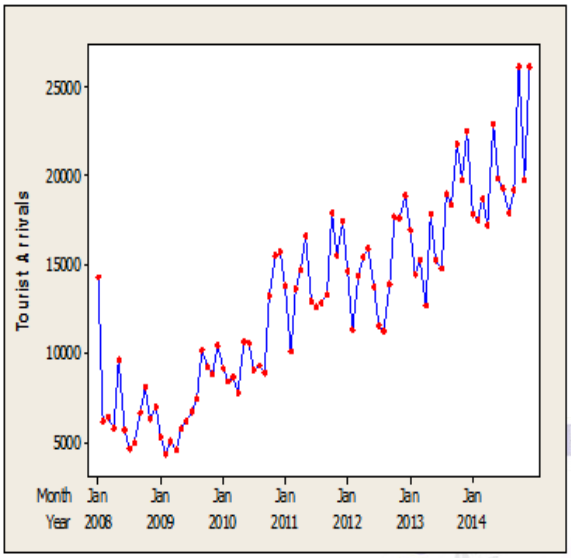

Figure 7: Time series plot of tourist arrivals Tourist arrivals from India.

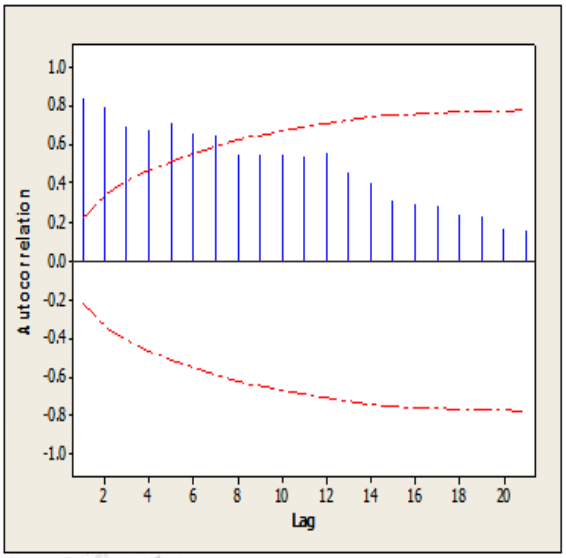

Figure 8: Autocorrelation function of from India.

Tourist arrival pattern of Maldives

Time series plot of arrivals from Maldives (Figure 9) shows an increasing trend after 2013. But fluctuation of the series is very high, increasing with time. It means arrivals from Maldives are not steady. Figure 10 is the ACF of tourist arrivals from Maldives. It shows that the series is not stationary. It suggests seasonal behavior.

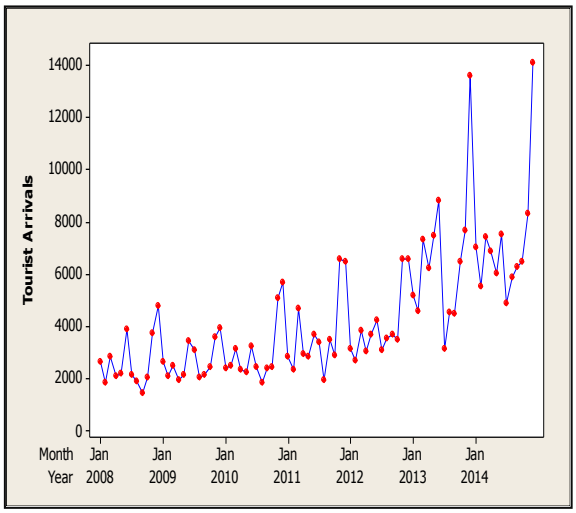

Figure 9: Time Series Plot of Tourist arrivals from Maldives

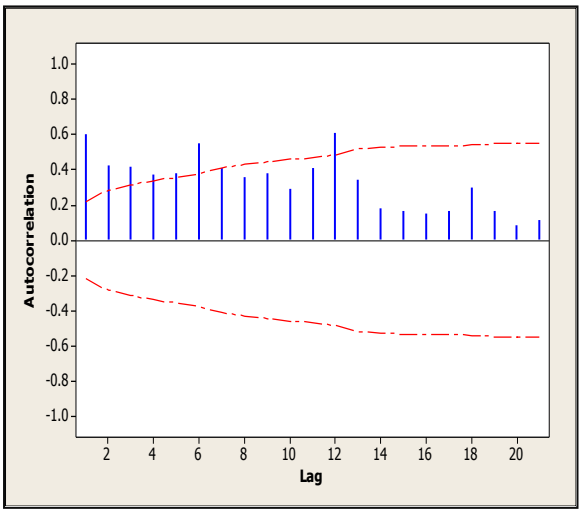

Figure 10: Autocorrelation function of tourist from Maldives

Tourist arrival pattern of China

Time series plot of arrivals from China (Figure 11) shows the arrivals are

constant up to the year 2012 November. After that, it shows high increasing of tourist arrivals. This movement is not steady. Figure 12 is the ACF of tourist arrivals from China. It is clear that the series is not stationary. But there is no seasonal pattern in the series.

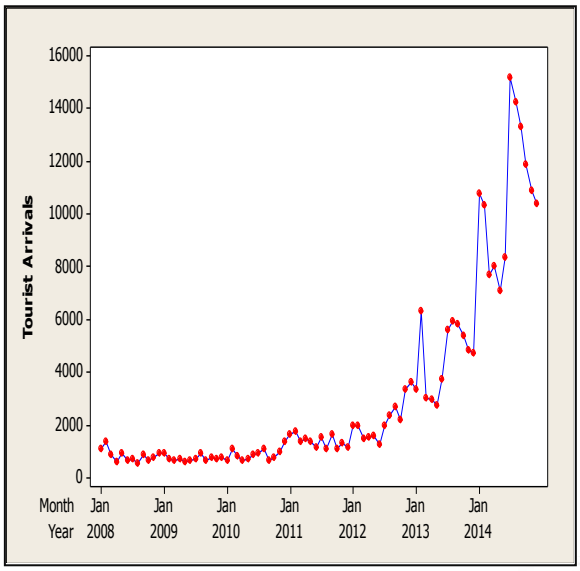

Figure 11: Time series plot of tourist from arrivals from China.

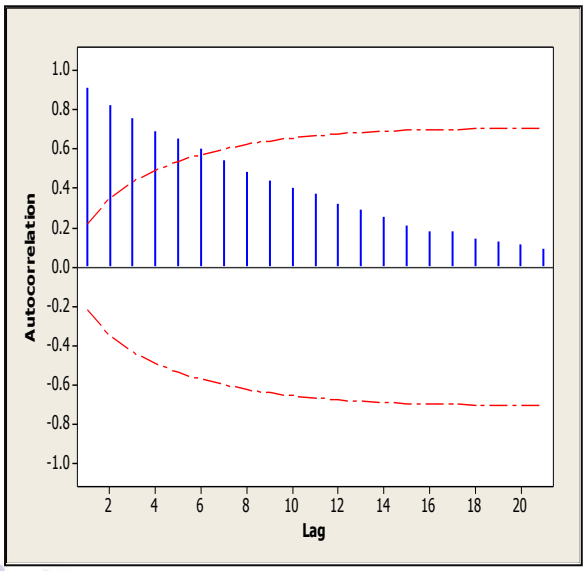

Figure 12: Autocorrelation Function for Tourist arrivals from China.

Tourist arrival pattern of Japan

Time series plot of arrivals from Japan (Figure 13) shows the arrivals are constant up to the year 2010 July. After that the arrivals shows heavy fluctuations, increasing with time. It clearly shows the increasing trend. This movement is not steady. Figure 14 is the ACF of tourist arrivals from Japan. It shows that the series is not stationary. But the seasonal pattern exists in the series.

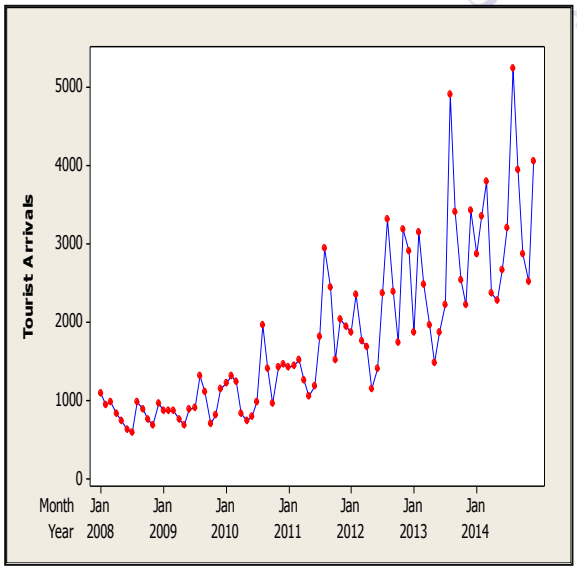

Figure 13: Time Series Plot of Tourist arrivals from Japan.

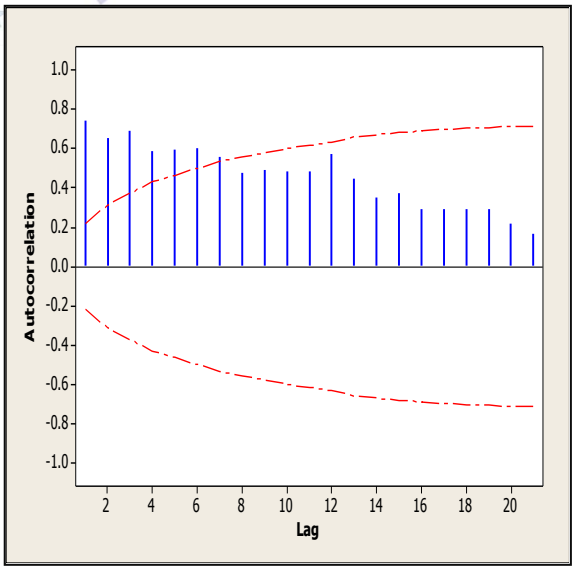

Figure 14: Autocorrelation function for Tourist arrivals from Japan.

Tourist arrival pattern of Pakistan

Time series plot of arrivals from Pakistan (Figure 15) shows the arrivals are decreased after 2008 April and slight increasing up to an end of 2010. This movement is not steady. Figure 16 is the

ACF of tourist arrivals from Pakistan. It shows that the series is not stationary. But there is a seasonal pattern suggest from the series.

Figure 17 shows the comparison of tourist arrivals from all countries in Asia. It is clear that India is on top and Pakistan at the bottom. But, arrivals from Maldives and China were higher than Japan and Pakistan. Figure 17 also confirm the increasing trend of all the countries. There are a significant growth and difference of tourist arrivals from India, compared to other countries in the Asian region.

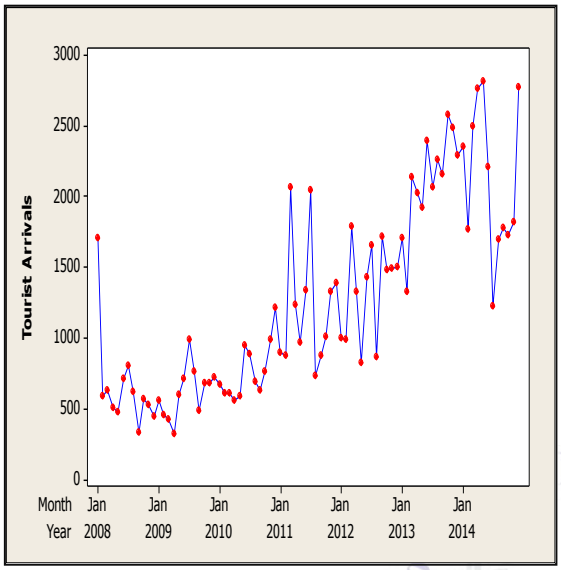

Figure 15: Time series plot of tourist arrivals Tourist arrivals from Pakistan

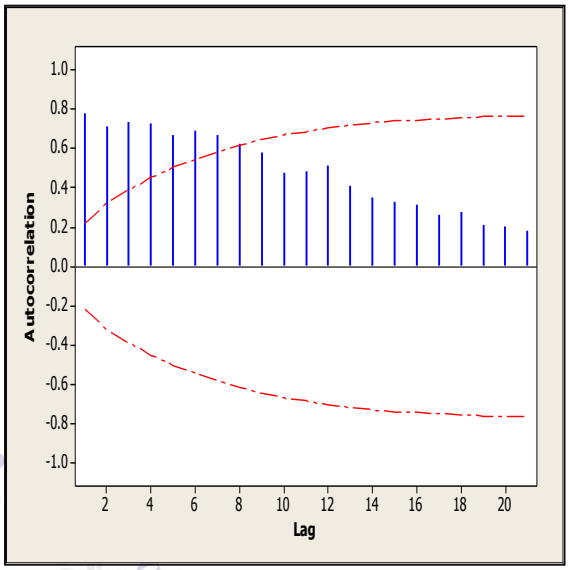

Figure 16: Autocorrelation function of Tourist arrivals from Pakistan

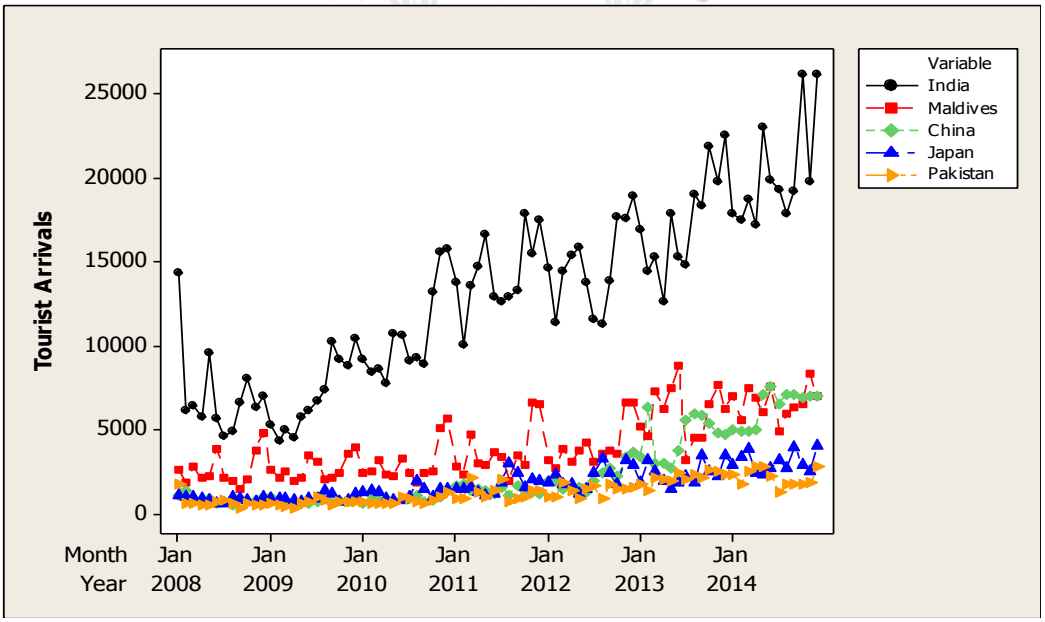

Figure 17: Time Series Plot of Tourist arrivals from all countries from Asia

Comparison of tourist arrivals from all countries

It was intended to test whether an average number of arrivals from different countries are same or not. One way - Analysis of Variance (ANOVA) was used for the purpose as follows;

H_0: $\mu_{India} = \mu_{Maldives} = \mu_{China} = \mu_{Japan} = \mu_{Pakistan}$

H_1 : *At least, one mean is different from the others*

P value of the ANOVA is (0.000) was significant at 5% significance level. Therefore, it was concluded that at least one mean is different from others. In other words, arrivals from at least one country are different from the others. Individual 95% Confidence Interval (CI) for means was obtained, given in Table 1.

Table 1: ANOVA

Region	N	Mean	StDev	Individual 95% CIs For Mean Based on Pooled StDev --------+---------+---------+---------+-
India	84	13036	5369	(*-)
Maldives	84	4074	1907	(-*)
China	84	2394	2153	(-*-)
Japan	84	1751	909	(-*-)
Pakistan	84	1261	693	(-*)
Pooled Standard Deviation			2771	--------+---------+---------+---------+- 3500 7000 10500 14000

Excluding India and Maldives, CI's of all the other countries overlap. It means there is no significant difference among arrivals from China, Japan, and Pakistan. Tourist arrivals from those countries were below 2500. India is producing 13000 tourists to Sri Lanka per month and Maldives producing 4000 tourists.

DISCUSSION

Series of arrivals were not stationary type. As such, Auto Regressive Integrated Moving Average (ARIMA) methods may not be suitable in forecasting arrivals. But Seasonal ARIMA technique may appropriate for the Asian region. It is recommended to test Moving Average Methods, Exponential Smoothing techniques, Holt's Winter's three parameter models, Decomposition techniques, linear and non-linear trend models for forecasting arrivals. Arrivals from all the countries show wave-like patterns. The "Circular Model (CM)" is a newly developed univariate time series forecasting technique, suitable for modeling wave-like patterns. [15] It is recommended to test the CM on arrivals.

CONCLUSION

The purpose of this study was to identify time series patterns of tourist arrivals from top five Asian countries to Sri Lanka. Patterns of tourist arrival from Asian countries to Sri Lanka show increasing trends. Arrival series of all the five countries were non-stationary. There was no significant autocorrelation in tourist arrivals. It is similar to the results of pattern recognition of tourist arrivals from all regions. [13-2] It means there is no significant correlation between present arrivals and the past. A significant difference of tourist arrival can be seen from India compared with other Asian countries.

REFERENCES

1. Konarasinghe, KMUB. International Tourism Demand on Attraction of Sri Lanka .In: Conference Proceedings.1st Wayamba International Research Conference; 2014, Wayamba University of Sri Lanka; p.75.
2. Konarasinghe, KMUB, Deheragoda, CKM. Modeling International Tourism Demand in Sri Lanka: Parametric Approach; In: Conference Proceedings. 2nd International Research Conference on Humanities and Social Sciences; 2013, Faculty of Humanities and Social Sciences University of Sri Jayawardanapura, Sri Lanka; p.118.
3. Sri Lanka Tourism Development Authority (SLTDA). (2014).Summary Report [Internet] 2014, [updated 2014 Dec; cited 2015 April] Available from http://www.sltda.lk/statistics.
4. Konarasinghe, WGS, Abeynayake, NR, Gunaratne, LHP. Forecasting Share Prices of Banks of Colombo Stock Exchange; In: International Journal of Research & Review; 2015. 2(6), p. 372-378.
5. Bemdt, D, Clifford, J. Using Dynamic Time Warping to Find Patterns in Time

Series; In: AAAI Technical Report WS-94-0; 1994.
6. Singh, SA. Long Memory Pattern Modelling and Recognition System for Financial Forecasting; In: Pattern Analysis and Applications; 1999. 2(3). p. 264-273.
7. Sternickel, K. Automatic pattern recognition in ECG time series; In: Computer Methods, and Programs in Biomedicine; 2002.p.109-115.
8. Morales-Esteban, A, Martínez-Álvarez,F, Troncoso, A, et.al. Pattern recognition to forecast seismic time series; In: Expert Systems with Applications; 2010.p. 8333-8342.
9. Gaebler, J., Lommatzsch,A., De Luca,E., & et.al. Pattern Recognition and Classification for Multivariate Time Series; 2011.
10. Shinde, S.P. & Deshmukh,V.P. Implementation of Pattern Recognition Techniques and Overview of its applications in various areas of Artificial Intelligence; In: International Journal of Advances in Engineering & Technology; 2011.1(4). p. 27-137.
11. Lin,J., Williamson,S., Borne, K., & et.al. Pattern Recognition in Time Series; 2012.
12. Konarasinghe, W .G. S. & Abeynayake, N. R. Time-series patterns of Sri Lankan stock returns; In : Proceedings of the 11th International Conference on Business Management, Ph.D. Consortium, Faculty of management studies and commerce , University of Sri Jayewardenepura; 2014. p. 78-95.
13. Konarasinge, K.M U.B. Time Series Patterns of Tourist Arrivals to Sri Lanka; In: Conference Proceedings of Interdisciplinary Business and Economics Research 2016, Bangkok; 2016.
14. Sharma, P., & Kaur, M. Classification in Pattern Recognition; In: A Review. International Journal of Advanced Research in Computer Science and Software Engineering; 2013.3(4). p. 298-306.
15. Konarasinghe, W.G.S., Abeynayake, N.R., & Gunaratne, L.H.P. Circular Model on Forecasting Returns of Sri Lankan Share Market; In: International Journal of Novel Research in Physics, Chemistry, and Mathematics, 2016. 3(1).p. 49-56. [Online] Retrieved from website: www.noveltyjournals.com

How to cite this article: Konarasinghe KMUB. Patterns of tourist arrivals to Sri Lanka from Asian countries. International Journal of Research and Review. 2016; 3(11):69-79.

******