

Comparison of Forecasting ability of Sama Circular Model, ARIMA and Decomposition Techniques

Dr. W. G. Samanthi Konarasinghe Institute of Mathematics and Management, Sri Lanka

samanthi@imathm.edu.lk

ABSTRACT

Time series data comprises several components; Trend, Seasonal variations, cyclical variations and irregular variations. These series follow irregular wave like patterns. This type of data is common in the fields of, Meteorology, Agriculture, Share Market Economic, Education, Healthcare and more. The Decomposition technique and the Auto Regressive Integrated Moving Average (ARIMA)/Seasonal Auto Regressive Integrated Moving Average (SARIMA) are the widely applied methods for forecasting such a series. Yet these techniques are unable to model the cyclical variation and they have some other weaknesses. According to the literature, modeling cyclical variation is highly important and crucial. Some researchers have attempted the Artificial Neural Network for the purpose, yet the success of them were doubtful. There were no Statistical techniques for the purpose. The Sama Circular Model (SCM) is a recently joined member to the family of forecasting techniques, developed on Newton's law of Circular Motion, Fourier transformation and Least Square Estimation. Indeed it is a frequency domain model. The SCM is capable in capturing all the components of a time series; Trend, Seasonal and Cyclical. Hence it was intended to compare the forecasting ability of the SCM with Decomposition techniques and ARIMA/SARIMA. Monthly female unemployment rates of Australia for forty years from 1978 were used for model testing. Goodness of fit tests and measurements of errors were used in model validation and verification. The SCM was superior to the ARIMA/SARIMA and Decomposition techniques.

Keywords: Sama Circular Model (SCM), Decomposition technique, ARIMA/SARIMA

1. INTRODUCTION

1.1 Background of the Study

Predicting the future was one of the strongest desires of a man. The history of predictions is linked with Babylonians and ancient Greeks. For thousands years ago, people in ancient Greece consulted the Oracle at Delphi to see the future of them (Daniel, 2013). Astrology was originated in Babylon around 2,400 years ago and wide spread all over the world. Those predictions were important in various ways. The society was interested in predicting weather conditions, as it is a guide line for seeking food and shelter, agriculture, trade and travel; Kings wanted to get the help for affairs of state and so on. Those non-scientific predictions are said to be subjective and fatalistic, yet they were the foundation for scientific forecasting.

Scientific prediction or Forecasting relies on mathematical modeling. A mathematical model is a simplification of a real world situation into an equation or a set of equations. They are easy to use and cost effective. The requirement of model for mathematical completeness can provide a conceptual framework which may help pinpoint areas where knowledge is lacking, and might stimulate new ideas and experimental approaches. It provides strategic and tactical support to a research motivating scientist.

All mathematical models have parameters; they are estimated in different ways. In an empirical model, parameters do not have any biological interpretation, but in mechanistic models, parameters have specific biological interpretation. Therefore empirical models are known as "Black- Box" models and "mechanistic models are known as "White- Box" models. A model is said to be "Static" when it does not have time- dependent component. In contrast, dynamic models contain time-dependent component. Deterministic models are not associated with any randomness. Conversely, in a stochastic model, randomness is present and variable states are described by associated probability distributions. In general stochastic models are referred as Statistical models (Konarasinghe, 2016). One of the most powerful and applied methodologies for generating forecasts is time series analysis. A data set containing observations of a single variable over a period of equal time intervals is called a time series. These time intervals can be years, months, days, hours, minutes and so on. In time series data, both values and the ordering of the data points have meaning (Stephen, 1998). Time series analysis and its applications have become increasingly important in various fields of research, such as business, economics, agriculture, engineering, medicine, social sciences, politics etc.

1.2 Research Problem

A time series has four components. They are; Trend, Seasonal, Cyclical and Irregular component. Hence these data follows irregular wave like patterns with trends. The Decomposition techniques and Auto Regressive Integrated Moving Average (ARIMA)/ Seasonal Auto Regressive Integrated Moving Average (SARIMA) methods and are the widely applied techniques in forecasting such series. The Sama Circular Model (SCM) is a recent development, in year 2018 for the purpose. The Decomposition technique is applied in several steps, first fit the trend model to the series and de-trend the series; secondly estimate the seasonal components to the de-trended series and then de-

sesonalized it. This de-sesonalised series is taken as the Irregular component, assuming the cyclical component is not present in the time series. This assumption is the main weakness of the technique. The ARIMA methodology explained present value of a series (Y) by past values of Y itself and the stochastic error terms. Although the ARIMA method can handle data with a trend, it does not support time series with a seasonal component. The Seasonal Auto Regressive Integrated Moving Average (SARIMA) is an extension to ARIMA that models the seasonal component of a series as well. Yet the ARIMA / SARIMA techniques have few weaknesses. Studies of; Ayodele et. al. (2014), Konarasnghe (2016, 2019), Konarasinghe (2020) and many others has shown that the ARIMA/ SARIMA forecasts do not follow the pattern of actual series. Also, they are applicable only if the data series has constant mean and constant variance (stationary series). On the other hand, SARIMA is unable to separate the seasonal and cyclical variations of a series. Yet, the Sama Circular Model (SCM) is applicable in modeling non stationary series, and separates seasonal and cyclical variations of a series. However the SCM is a newly developed model, which has not been much applied in real life situations. Hence it is worth testing the SCM on more real life data and compares the forecasting ability of SCM with the well known Decomposition techniques, ARIMA/ SARIMA methods.

1.3 Objective of the Study

To compare the forecasting ability of the SCM, Decomposition techniques and ARIMA/ SARIMA techniques

1.4 Significance of the Study

A wave like pattern is common in real life time series data sets. For examples, share market prices, tourist arrivals, market prices of goods, weather data (temperature, wind, rainfall, runoff etc.); unemployment; consumer price index; human blood sugar or blood pressure levels; harvest of crops; spread of a disease and so on. This pattern may contain trend, seasonal, cyclical and irregular variations. A trend is the overall tendency of the data series, either upward or downward. A seasonal variation is a wave like pattern with period of oscillation less than 12 months whilst a cyclical variation is a wave like pattern with period of oscillation greater or equals to 12 months. Accurate and reliable trend analysis is possible with Trend models; Linear, Quadratic, Growth Curve etc. The decomposition technique and SARIMA can be used to model seasonal variation but not to capture the cyclical variation. Also, a series may contain more than one seasonal variation, yet these techniques are unable to capture them.

Modeling seasonal and cyclical patterns are extremely important in various ways. For examples, identifying seasonal patterns of rainfall data is useful in, Agriculture; Tourism; Transportation; Healthcare; Business and all the day to day activities, identifying seasonal patterns of runoff is essential for flood forecasting, disaster management etc. Seasonal patterns related to sales or number of customers to a particular restaurant helps to decide the workforce requirements and variety of foods needed. Identification of cyclical variations is crucial in Business, Meteorology and Pandemics etc for strategic

decisions. As per Helena, et al., (2010) business cycles have been a major topic of discussion, especially in periods of large fluctuations in aggregate variables of industrialized economies to understand their consequences to the economy. Frequently some part of the world suffers from cyclical climatic changes like Tsunami, Extreme floods, Volcanoes etc. At present the entire world suffers from a pandemic which confirms a cyclical pattern.

The SCM is captures Trend, Seasonal and Cyclical variations easily and differentiate among several seasonal or cyclical patterns within a series. Yet it is a recently developed technique, needed more applications to confirm the applicability of the method. This study will provide a better understanding about the forecasting ability of the SCM in comparison to the well known Decomposition techniques and ARIMA/SARIMA.

2. LITERATURE REVIEW

The literature review is based on the applications of the Decomposition Techniques, ARIMA/ SARIMA and the Sama Circular Model.

The study of Sen & Chaudhuri (2016) was focused on testing the applicability of the decomposition techniques in forecasting stock market indices of the Auto sector of the Indian stock market. The data collection period was January 2010 to December 2015. The authors have estimated the Trend, Seasonal and Irregular components of the series and concluded that the decomposition technique is suitable for the purpose. Yet they have not reported about the satisfaction of model validation and verification criterion. Prem & Rao (2015) have tested both Additive and Multiplicative Decomposition models for forecasting wind speed. They have used the wind speed data collected at every minute for a period of 5 months at a wind farm in Bagalkot, Karnataka, India. They have compared the forecasting ability of the Decomposition models with Exponential Smoothing Techniques, ARIMA and some other Time Series by means of Mean Absolute percentage Error (MAPE) and Root mean Square Error (RMSE) and concluded that the Decomposition models are superior to others. Konarasinghe (2016) intended to forecast tourist arrivals to Sri Lanka from the Western European Region. Author has tested both additive and multiplicative models for five countries of the region, by using monthly data for the period from January 2008 to December 2014. Author has concluded that the Additive Decomposition models are the best for the purpose. None of the studies; Sen & Chaudhuri (2016), Prem & Rao (2015), Konarasinghe (2016) has attempted to estimate the possible cyclical component of the series.

The study of Prapanna, et al., (2014) was focused on applying ARIMA on forecasting Indian stock market. Monthly average share prices for eight companies of the National Stock Exchange for the period from April 2012 to February 2014 were used for the analysis. The Akaike Information Criteria (AIC) was used to quantify the goodness of fit of the model. The Mean Absolute Error (MAE) was used to assess the forecasting ability of the selected models. Authors have recommended the ARIMA in forecasting share prices of the individual companies of the Indian stock market. Emenike (2010) has tested ARIMA models to forecast stock returns of the Nigerian Stock Exchange (NSE).

Monthly All-Share Indices of the NSE from January 1985 to December 2009 were used for the model fitting and validation. The ARIMA (1, 1, 1) was selected as the best fitting model, but forecasted returns were not match with actual returns. Konarasinghe (2016) has tested the ARIMA and the Circular Model (CM) on forecasting Sri Lankan share market. The sample of the data analysis consist fifty randomly selected companies listed in the Colombo Stock Exchange. Results revealed that both ARIMA and the CM are successful in forecasting individual company returns, but the ARIMA forecasts did not follow the patterns of actual returns whilst the CM forecast did. The study of Konarasinghe (2018) is an application of SARIMA and Decomposition techniques to forecasting foreign guest nights in Southern Coast of Sri Lanka. The study revealed that the SARIMA is superior to the Decomposition techniques for the purpose.

The study of Konarasinghe (2018) was focused on testing the SCM, SARIMA and the Decomposition techniques on forecasting tourist arrivals to Sri Lanka. The author has utilized monthly data from April 2009 to December 2016. Results revealed that the Decomposition Multiplicative Model and SCM served the purpose with satisfactorily low measurements of errors, but the SARIMA did not. Hence the forecasting ability of Decomposition Multiplicative Models and SCM were compared and found that the SCM is superior to Decomposition method in light of that. Another study of Konarasinghe (2018) attempted to use SCM on forecasting two main stock market indices of Sri Lanka; the All Share price Index (ASPI)and the S&P SL 20. Monthly ASPI from January 2009 to March 2018 and monthly S&P SL 20 indices from June 2012 to March 2018 were used for the analysis. It was concluded that the SCM is a suitable model for forecasting Sri Lankan stock market indices. Konarasinghe (2019) has tested the SCM and ARIMA on forecasting BSE Sensex index of the Indian stock market, using monthly data for the period from January 1980 to April 2019. Author has recommended both the techniques for the purpose based on low the measurements of errors, yet mentioned that the pattern of ARIMA forecast not aligned with the actual values. Konarasinghe (2020) has tested the SARIMA and SCM on Forecasting Foreign Guest Nights in Anuradhapura of Sri Lanka. Monthly data for the period of January 2008 to December 2017 obtained from the Sri Lanka Tourism Development Authority (SLTDA) were used for the model based analysis. Results of the study revealed that the SCM is suitable for the purpose, but not the SARIMA.

As per the review, Decomposition techniques, ARIMA/ SARIMA and SCM were successful in various fields of studies. Yet, the SCM does not have the weakness of the Decomposition technique and ARIMA/ SARIMA.

3. METHODOLOGY

The study is based on the Decomposition techniques, ARIMA/ SARIMA models and Sama Circular Model (SCM) of Konarasinghe (2018). Monthly female unemployment rates of Australia from 1978 to 2018 were obtained from the Australian Bureau of Statistics official website for model testing. Time Series plots and Auto Correlation Functions (ACF) were used for pattern recognition. Then the three models were tested,

using statistical software Minitab 17. Goodness of fit tests and measurements of errors were used in model validation. The ACF of residuals and Ljung-Box Q statistics (LBQ) were used to test the independence of residuals. The Probability plot and the Anderson Darling test were used to test the normality of residuals. Forecasting ability of models was assessed by Mean Square Error (MSE) and Mean Absolute Deviation (MAD).

3.1 Decomposition Techniques

The Decomposition technique splits a time series into four components; Trend (T), Cyclical influences (C), Seasonal influences(S) and the random error (I). Two general types of decomposition models are Additive and Multiplicative models. The multiplicative model is given as;

$$Y = T.S.C.I \tag{1}$$

The additive model uses when the size of the seasonal pattern does not depend on the level of the data. Model is;

$$Y = T + S + C + J \tag{2}$$

3.2 Auto Regressive Integrated Moving Average (ARIMA)

An ARIMA model is given by:

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \tag{3}$$

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

$$\theta(B) - 1 - \theta_1 B - \theta_2 B^2 \dots \theta_p B^q$$

 $\varepsilon_t =$ Error term, D = Differencing term,

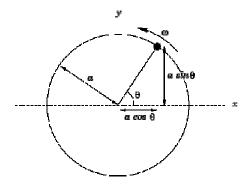
B = Backshift operator ($B^{a}Y_{t} = Y_{t-a}$)

3.3 Sama Circular Model (SCM)

The SCM is the improved form of the Circular Model (CM) of Konarasinghe (2016). Development of the SCM was based on the Fourier transformation (FT), Least Square Regression and Differencing technique.

A particle *P* moving in a horizontal circle of centre O and radius *a is* given in Figure 1. The ω is the angular speed of the particle;

Figure 1: Motion of a Particle in a Horizontal Circle



At one complete circle $\theta = 2\pi$ radians. Therefore, the time taken for one complete circle (*T*) is given by: $T = 2\pi / \omega$ (4)

In circular motion, the time taken for one complete circle is known as the period of oscillation. In other words, the period of oscillation is equal to the time between two peaks or troughs of sine or cosine function. If a time series follows a wave with f peaks in N observations, its period of oscillation can be given as;

$$T = \frac{\text{total number of periods}}{\text{total number of peaks}} = \frac{N}{f}$$
(5)

Equating (4) and (5); $\frac{2\pi}{\omega} = \frac{N}{f}$

Hence for a random variable Y_t , the SCM is written as;

$$(1-B)^{d} Y_{t} = \sum_{k=1}^{n} (a_{k} \sin k\omega t + b_{k} \cos k\omega t) + \varepsilon_{t}$$
(6)

Where; d^{th} order difference of Y_t , $Y_t^d = (1 - B)^d Y_t$, B is the Back Shift operator.

Model assumptions are;

 Y_t is a continuous random variable $t \ge 0$, $k \in R^+$ Series, sin *kot* and cos *kot* are independent ε is Normally distributed and independent

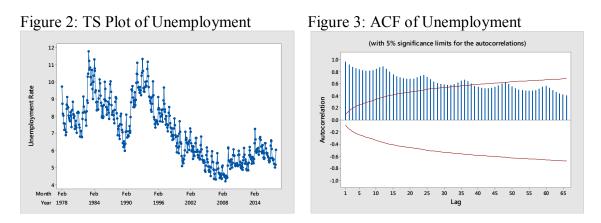
4. **RESULTS**

The analysis of the study contains five main parts;

- 4.1 Pattern recognition4.2 Testing the Decomposition techniques4.3 Testing the ARIMA4.4 Testing the SCM
- 4.5 Model comparison

4.1 Pattern Recognition

Figure 2 is the Time Series (TS) plot and Figure 3 is the ACF of the female unemployment of Australia;



The time series plot shows heavy fluctuations with overall decreasing trend. The ACF confirms the trend and non stationary of the series. In other words, the series follows an irregular wave like pattern with trend.

4.2 Testing the Decomposition Techniques

The Decomposition Multiplicative Model and Additive Model were tested and found that the residuals of both models were not normality distributed or independent.

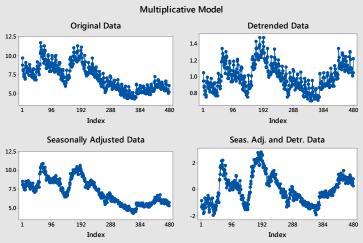


Figure 4: Component Analysis-Multiplicative Model



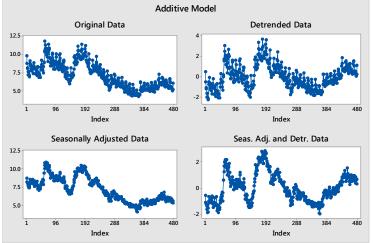
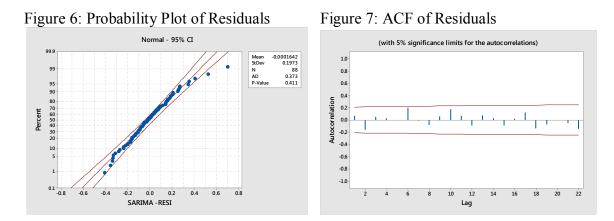


Figure 4 and Figure 5 are the component analysis of Multiplicative model and Additive model respectively. The de-trended and seasonally adjusted data series of both graphs show clear patterns. Hence the Decomposition technique cannot be used to forecast the unemployment.

4.3 Testing the ARIMA

Figure 3 suggests the Seasonal ARIMA model. The ARIMA (0, 1, 1) $(1, 0, 1)_{12}$ satisfied the model validation criterion, normality and independence. Figure 6 is the probability plot of residuals and Figure 7 is the ACF of residuals;



Measurements of errors of the model were satisfactorily small. Hence the SARIMA model is suitable for the purpose. As per the fitted model, unemployment follows seasonal pattern with 12 months length.

4.4 Testing the SCM

The SCM was tested on the first difference series and the second difference series of the data set. The angular speed ω was calculated by $\omega = 2\pi f / N$; where f is the number of peaks and N is the number of observations in the series. For 0 < k < 7, 12 trigonometric series for $\sin k\omega t$ and 12 trigonometric series for $\cos k\omega t$ were obtained. The correlation analysis confirmed the independence of the trigonometric series. The best fitting SCM model is;

$$Y_{t} = Y_{t-1} - 2Y_{t-2} + 0.0066 + 0.1045 \sin 2\omega t + 0.3152 \cos 0.5\omega t + 0.4728 \cos \omega t$$
(7)

Residuals were normally distributed and independent. Measurements of errors of the model were satisfactorily small. Hence the SCM is suitable for the purpose.

The fitted SCM comprises 3 trigonometric functions; $\sin 2\omega t, \cos 0.5\omega t, \cos \omega t$, given in Figures 8, 9 and 10. Period of oscillation of Figure 8 is 6 months. It means unemployment follows seasonality in six months period with lowest in April and highest in June and so on. In Figure 9, period of oscillation is 19 months. Therefore unemployment follows a cyclical variation in 19 months time intervals; lowest in August and highest in February and so on. Also, Figure 10 shows a seasonal variation in 10 months intervals. Hence female unemployment of Australia has significant seasonal variations as well as cyclical variation.



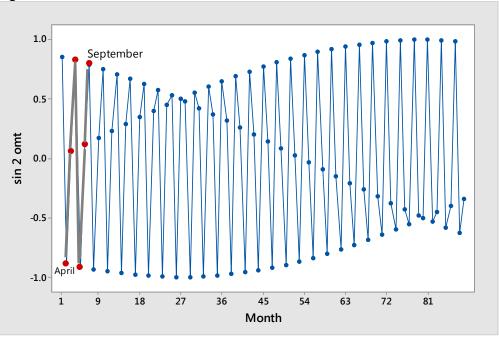
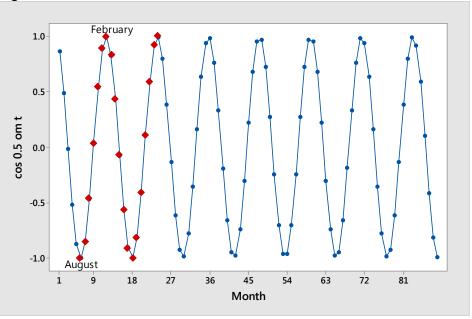
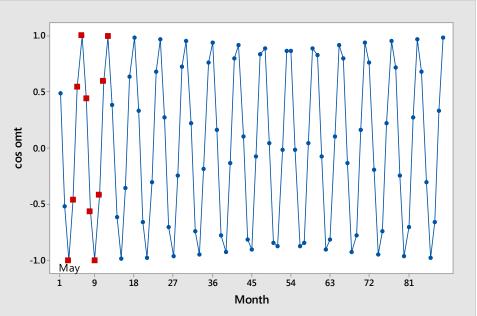


Figure 9: Plot of cos0.5ωt







4.5 Model Comparison

Model summary of SARIMA and SCM are given in Table 1;

Table	1:	Model	Summary
-------	----	-------	---------

Model	Model Fitting		Model	
			Verification	
$Y_{t} = Y_{t-1} - 2Y_{t-2} + 0.006 + 0.105\sin 2\omega t$	MAD	0.260	MAD	0.170
$+0.315\cos\left(\frac{1}{2}\right)\omega t + 0.473\cos\omega t$	MSE	0.114	MSE	0.035
ARIMA (0, 1, 1)(1, 0, 1) ₁₂	MAD	0.151	MAD	0.420
	MSE	0.038	MSE	0.249

Measurements of errors of SARIMA and SCM are almost equally small. Figure 11 shows the Actual data Vs. SARIMA forecasts and SCM forecasts. It is clear that the SARIMA forecasts are little over estimations, but SCM forecasts are much closer to the actual values. Also the pattern of SCM forecasts is aligned with the actual values.

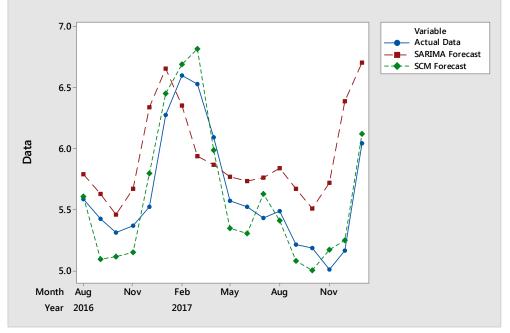


Figure 11: Actual Data, SARIMA Forecast and SCM Forecast

The SARIMA shows that the unemployment follows a seasonal behavior in 12 months intervals, but the SCM shows that the unemployment follows two seasonal patterns and a cyclical pattern. Accordingly the SCM gives more information than the SARIMA. Hence the SCM is superior to SARIMA in forecasting female unemployment of Australia.

5. CONCLUSION AND RECOMMANDATIONS

The study was focused on comparison of the forecasting ability of newly developed Sama Circular Model (SCM) with the well known Decomposition technique and the ARIMA/ SARIMA models. Female unemployment rates of Australia for the last forty years were used for the model testing and found that the SARIMA and SCM are suitable for the purpose. Also it is concluded that the SCM is superior to SARIMA in light of that.

The Decomposition techniques and SARIMA were highly used in forecasting wave like patterns of real life time series. Yet they have some weaknesses compared to the SCM. In both methods, it is necessary to specify the seasonal length in model testing; these techniques are unable to capture the cyclical variation of a time series. Most important property of the SCM is that it is able to separate the seasonal variations and cyclical variations without any effort. Hence it is recommended to apply the SCM in more real life situations of, Meteorology; Agriculture; Healthcare and Biological Sciences; Business, Finance and more to get the benefit of it.

REFERENCES

- Ayodele A. A., Aderemi O. A., & Charles K. A.,(2014). Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. *Journal of Applied Mathematics*, 1-6. http://dx.doi.org/10.1155/2014/614342
- Daniel C. S., (2013). The Future: A History of Prediction from the Archives of Scientific American. https://www.scientificamerican.com/article/50-100-the-future-history-predictionfrom-archives-scientific-american/
- Emenike, K.O. (2010). Forecasting Nigerian Stock Exchange Returns: Evidence from autoregressive Integrated Moving Average (ARIMA) Model. Department of Banking and Finance, University of Nigeria, Enugu Campus, Enugu State, Nigeria.

http://ssrn.com/abstract=1633006/

- Konarasinghe, W.G.S., (2016). *Model Development for Stock Returns*. Doctor of Philosophy Thesis, University of Peradeniya, Sri Lanka
- Konarasinghe, K.M.U.B., (2016). Decomposition Techniques on Forecasting Tourist Arrivals from Western European Countries to Sri Lanka. Conference Proceedings of the 13th International Research Conference on Business Management (ICBM), University of Sri Jayewardenepura, Sri Lanka. <u>https://ssrn.com/abstract=2910291</u>
- Konarasinghe, W.G.S., (2018). Sama Circular Model on Forecasting Sri Lankan Stock Market Indices. Mathematical Sciences International Research Journal. 7 (Special issue), 1-4.
- Konarasinghe, W.G.S., (2018). Sama Circular Model on Forecasting Tourist Arrivals to Sri Lanka. Mathematical Sciences International Research Journal. 7(2), 191-193.
- Konarasinghe, K.M.U.B., (2018). Forecasting Foreign Guest Nights in Southern Coast of Sri Lanka: An Application of SARIMA Model. International Journal of Novel Research in Physics Chemistry & Mathematics, 5(3):22-30.
- Konarasinghe, W. G. S., (2019). Sama Circular Model and ARIMA on Forecasting BSE Sensex. International Journal of Novel Research in Physics Chemistry & Mathematics. 6(2), 1-7. <u>https://www.noveltyjournals.com/issue/IJNRPCM/Issue-2-May-2019-August-</u>2019

- Konarasinghe, K.M.U.B.,(2020). Sama Circular Model on Forecasting Foreign Guest Nights in Anuradhapura of Sri Lanka. Review of Integrative Business and Economics Research, 9(3): 20-34.
- Helena, M., Ambrosio Dias, A., & Dias, J., (2010). Measuring the Cyclical Component of a Time Series: Proposal of a New Methodology. Brazilian Review of Econometrics, 30(1).
- Prapanna, M., Labani S., Saptarsi G. (2014). Study of Effectiveness of Time Series Modelling (ARIMA) in Forecasting Stock Prices. *International Journal of Computer Science, Engineering and Applications* (IJCSEA), .4(2), 13-29.
- Prema, V., K. & Rao, K., U., (2015). Time series decomposition model for accurate wind speed forecast. *Renewable: Wind, Water and Solar, a Springer Open Journal*. DOI 10.1186/s40807-015-0018-9

https://link.springer.com/content/pdf/10.1186/s40807-015-0018-9.pdf

- Sen., J., & Chaudhuri, T., D., (2016). Decomposition of Time Series Data of Stock Markets and its Implications for Prediction – An Application for the Indian Auto Sector. <u>https://www.researchgate.net/publication/290263627</u>
- Stephen, A.D., (1998). Forecasting Principles and Applications. First Edition. Irwin McGraw-Hill, USA.