Identifying Seasonality in Time Series by Applying Fast Fourier Transform

Hmeda Musbah, Mo El-Hawary, Hamed Aly
Department of Electrical and Computer Engineering, Dalhousie University, Halifax, Canada
Hm392855@dal.ca, elhawary@dal.ca, hamed.aly@dal.ca

Abstract— **The importance of studying time series is that most forecasting models assume that the time series must be stationary. In addition, non-stationary time series can cause unexpected behaviors or create a non-existing relationship between two variables. The aim of this study is to shine new light on the Fast Fourier Transform (FFT) technique through an examination of its efficiency in identifying the trend and seasonality by applying it to many time series. A comparison between the FFT technique and Autocorrelation Function (ACF) has been conducted as well. The results show that the FFT technique has acceptable performance in identifying the trend and seasonality. The most obvious observation is that, unlike the FFT technique, the ACF has limitations in determining the exact time of the seasonality that repeats itself.**

Keywords—Time Series, Seasonality, Fast Fourier Transform, Autocorrelation Function, Frequency Domain

I. INTRODUCTION

Several studies have been conducted on time series because of the importance in planning and decision-making; researchers were interested in studying and analyzing the time series and its trend, seasonal, cycle and random components. Understanding the behavior of these components helps build a successful model that can clarify the data to allow for prediction, monitoring, or control. Stationary time series is a term [1] that occurs when the mean, covariance and autocorrelation is constant along a time period of the time series. This is considered one of the most important terms that researchers are interested in studying because non-stationary time series are very hard to predict and can never be modeled or forecasted. The results gathered from the non-stationary time series can be false or misleading, in particular when a non-existing relationship between two variables exists. Also, some models and approaches, such as Box- Jenkins approach [1], assume in the first step that the time series should be stationary [2]. The trend component and the seasonal component are two of the components that could turn the time series into a non-stationary time series through two defining factors. These factors are when a varying mean occurs over time due to trends, and a changing variance due to seasonality. Therefore, many researches determined to detect whether there is trend component or seasonal component in the time series.

In this paper, Fast Fourier Transform Technique (FFT) has been applied to different data to evaluate the FFT's performance in detecting the seasonal component and compare it to Autocorrelation Function method (AFC).

II. LITERATURE REVIEW

In [3], four graphical techniques were used to detect the seasonality in the time series: a run sequence plot, a seasonal sub-series plot, multiple box plots and the autocorrelation plot. The authors stated that to apply seasonal subseries plot and box plot, the seasonal periods must be known. Likewise, [4] mentioned that the time plot of the whole series could be used to identify the seasonality by determining whether there are repeated peaks and troughs on regular periods and that have similar magnitude.

The Buys Ballot table has been used to detect the presence of the trend and seasonal effects in time series [5]. Several graphics that were proposed by the Buys Ballot table have been used to check the existence of seasonality [6]. Besides the graphical techniques, there are other statistical tests that could be used to identify the seasonality. These tests were summarized in [3] into three groups: the $\chi2$ Goodness-of-Fit test, the Kolmogorov-Smirnov goodness-of-fit test and the Nonparametric test. The authors applied different types of tests to the row variances of the Buys Ballot table where the Student *t*-test and Wilcoxon Signed-Ranks test showed good results in detection of seasonality. Wind speed in winter and summer have been compared to each other to study the seasonality by applying various signal processing techniques. The seasonality of wind speed in winter and summer time series were identified by Continuous Wavelet Transform (MCWT) [7].

The historical rainfall data at Ilorin North Central Nigeria was studied in [8], where they used Mann-Kendell trend analysis, Augmented Dickey Fuller (ADF) test, ACF to identify the trend component. The results showed that the observed data was non-stationary. Also, they used Fourier Transform to convert the historical rainfall data from the time domain to the frequency domain to detect the seasonality. The authors stated that the historical rainfall data had seasonality every 12, 6, and 4 months. One of the most common spread tools used to identify seasonality is the ACF tool. In [9] a Box-Jenkins algorithm was used for forecasting. This study focused on building a statistical model for forecasting the monthly average surface temperature in the BA region of Ghana to understand the dynamics of events. For determining the seasonality, authors used the visual inspection, decomposing the monthly average surface temperature into the various components and ACF. All these approaches confirmed the presence of seasonality. ACF plays a vital role in Box-Jenkins algorithm where it is considered the main tool to identify the trend and seasonal components. In [10]

978-1-7281-3406-2/19/$31.00 ©2019 IEEE

it was stated that the seasonality must be significant, otherwise the ACF tool cannot detect it.

III. Methods and Results

FFT plays a vital role in signal processing and data analysis. It computes the Discrete Fourier transform (DFT) of a signal and its inverse (IDFT), therefore, the FFT is a way to convert the time series from time domain to frequency domain. The number of computations in FFT technique is less than DFT, due to FFT technique reducing the complexity of computing by using the factor N/2 log N rather than N^2, where N is the number of points. The FFT technique is a useful algorithm because it can be used to identify the seasonality in the time series [8]. Hence, the FFT technique was applied to many sets of data to evaluate the efficiency of FFT technique in detecting the seasonality. See Table (1).

Figures (1a) and (2a) show mean daily temperature and monthly wind data that have been plotted as a time series respectively. The time series of mean daily temperature has a low frequency of regular pattern showing small rough edges throughout the sine wave. Also, the time series of monthly wind has a nearly smooth regular pattern up and down.

Table (1) shows many sets of data and its time period.

Time series	Time period	Reference
Mean daily temperature of Saugeen River	Jan 01, 1988 to Dec 31, 1991	[9]
Monthly wind of Halifax	Jan.2000 – Dec.2016	[9]
Hourly internet traffic data	06:57 hours on 7 June to 11:17 hours on 31 July 2005	[9]
Monthly Sutter county workforce	Jan.1946 – Dec.1966	[9]
Monthly Civilian labour force in Australia	Feb 1978 – Aug 1995	[9]
Monthly Nigeria power consumption		
Annual unemployment, U.S.	1890 to 1970	[9]

The most important concept from the output of the FFT technique that we focus on is the significant spikes and their locations. Clearly the time series is formed from one frequency. This frequency is represented by a significant spike which indicate that there is a seasonality and the time series tends to repeat its self every 365.33 days. Figure (1b) shows the FFT technique output of mean daily temperature. While figure (2b) shows the result of running FFT technique.

The significant spike that is located at $3.445 \times 10^{-8}\ Hz$ means the time series of wind speed tends to iterate itself every 12 months. Neither the times series of mean daily temperature nor the time series of the monthly wind need to apply the FFT technique or any other techniques, this is due to the clear presence of seasonality. The FFT technique is still required to determine the time that the series repeats itself.

Figure (3a) shows the time series of hourly internet traffic data where the seasonality protrudes noticeably. This is illustrated in the frequency domain as shown in Figure (3b) as the significant spike tends to repeat itself every 23.67 hours, as well as harmonics and noise components.

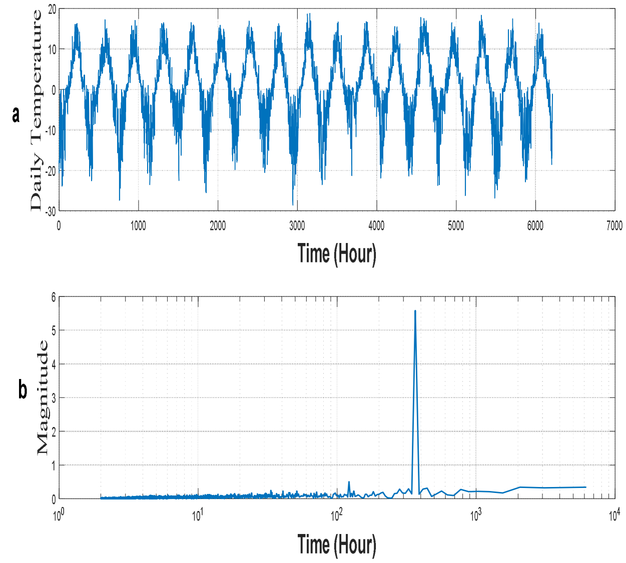

Figure (1) a) Shows the time series of the daily temperature data. b) Shows the FFT technique output of mean daily temperature.

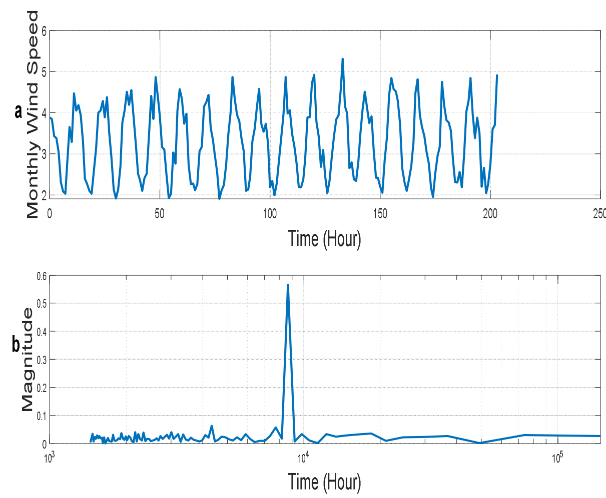

Figure (2) a) The time series of monthly wind data. b) The result of running FFT technique.

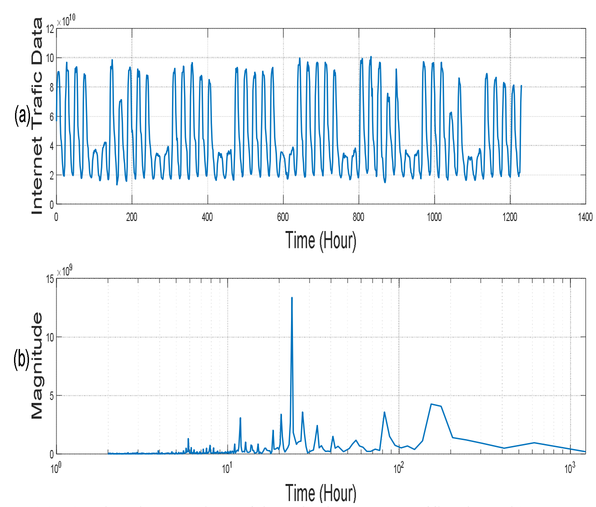

Figure (3) a) The time series of hourly internet traffic data. b) Presents the output of FFT technique for the time series of hourly internet traffic data.

Figures (4a) and (5a) display the time series of monthly Sutter County workforce and monthly civilian labour force respectively. Through the visual inspection, it could be confirmed that the two-time series have a significant upward trend where this significant trend could hinder detection of the seasonality. Especially if the seasonality was insignificant as shown in Figure (5a). Figure (4b) illustrates several spikes which are located at different times. The significant spike indicates that the time series of monthly Sutter County workforce has a seasonality which is repeated every 12 months. Sometimes it is difficult to verify if the seasonal component is included or not in a time series through either the visual inspection as shown in Figure (5a) or the FFT technique as shown in Figure (5b).

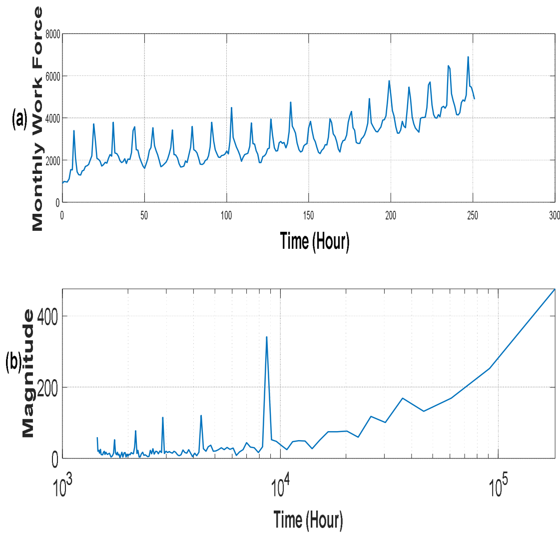

Figure (4) a) The time series of monthly Sutter County workforce. b) Illustrates the component that forms the time series of monthly Sutter County workforce.

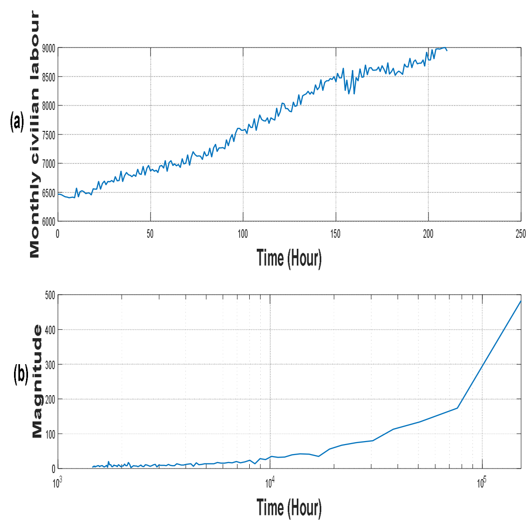

Figure (5) a) Monthly civilian labour force. **b)** Result of running FFT technique.

Figure (7a) shows that the swings in the time series of monthly Nigeria power consumption are trivial and due to the presence of a significant trend the FFT technique was not capable of recognizing the seasonality as shown in Figure (7b).

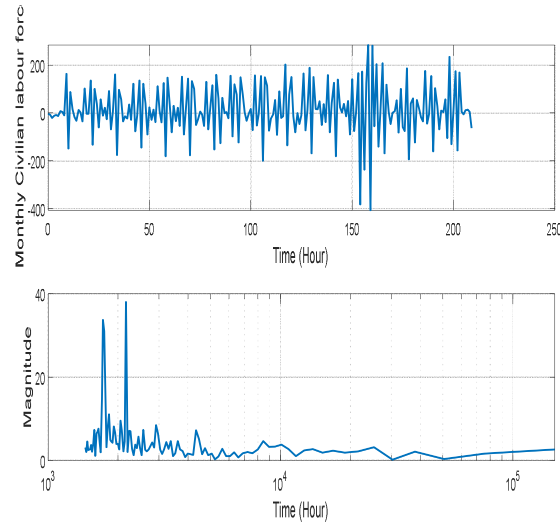

Figure (6) a) The time series of monthly civilian labour force after applying first differencing. b) Output of FFT technique for the time series of monthly civilian labour force after applying first differencing.

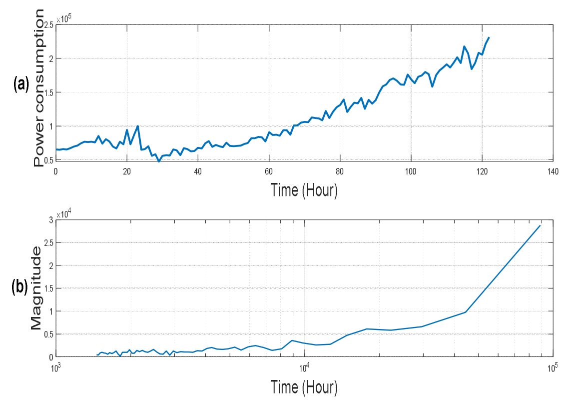

Figure (7) a) The time series of monthly Nigeria power consumption. b) Result of applying FFT technique to the time series of monthly Nigeria power consumption.

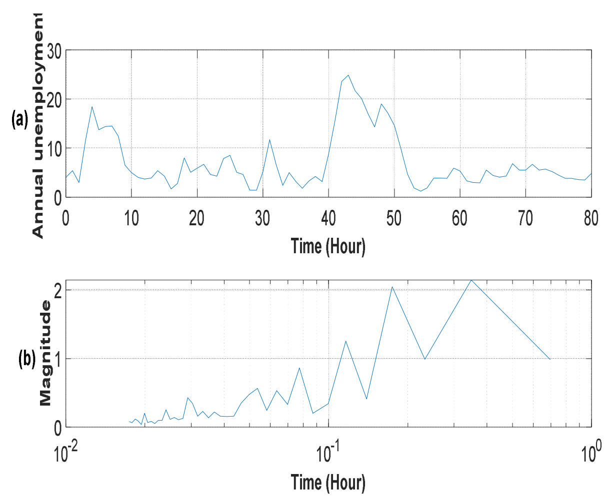

Figure (8) a) Time series of annual unemployment. b) Output of applying FFT technique to the time series of annual unemployment.

Figure (8a) shows time series of annual unemployment. Through visual inspection, it is clear the time series indicated has unpredictable behavior and it could be confirmed that there is no repetition pattern. The output of running FFT technique displays insignificant spikes as show in Figure (8b). This shows that the time series of annual unemployment does not include a seasonal component.

IV. Autocorrelation Function (ACF)

ACF measures the relation between every two consecutive observations in the time series. Several studies have used the ACF to determine whether the time series is stationary or not, and to identify the existence of the seasonality [8], [9], [10]. In this study, the FFT technique results have been validated by comparing its results to the ACF results. Figure (9) shows the output of ACF of the time series of monthly wind speed. The ACF plot illustrates a sine wave with minimal reduction in length as lag increases which results in tailing off. The spikes at lags 1, 12 and 24 for monthly data indicate that there is seasonality which repeats itself every 12 months. This result supports the FFT technique. Figure (10) shows the output of ACF of the time series of electrical load. The ACF has significant positive spikes at lags 24, 48 and 72 for hourly data. In comparing these results to the FFT results that are presented in [11] we found that the ACF shows us the time series of electrical load has a seasonality at 24 hours, while the FFT shows us there is seasonality at 24 and 12 hours.

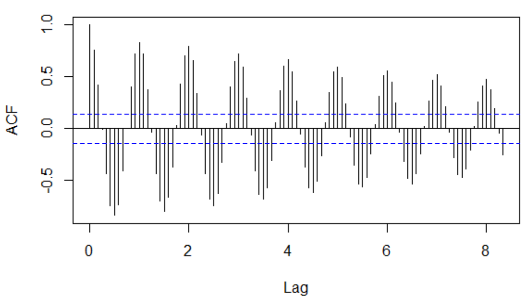

Figure (9) Output of ACF of the time series of monthly wind speed.

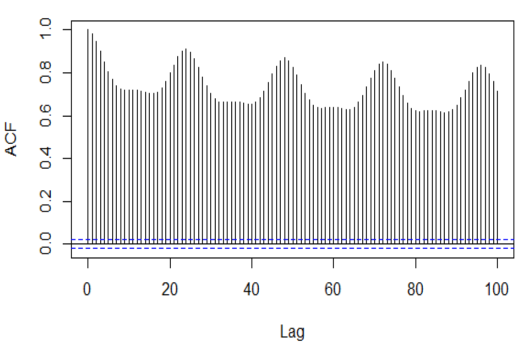

Figure (10) Output of ACF of the time series of electrical load.

V. Conclusion

Many sets of time series have been analyzed by applying the FFT technique to determine the seasonal component. The results of the FFT technique displayed a good performance in detecting the seasonality, but there are limitations that prevent the FFT technique from identifying the seasonality when the series contains a significant trend or insignificant swing along the periods of the time series. The output of the FFT technique showed good results in determining the trend. Correspondingly, the FFT technique outperformed the ACF. The ACF has restrictions in identifying the exact time of the seasonality that is repeated. The seasonality must be significant before the ACF is able to detect it.

References

[1] R. Manuca and R. Savit, "Stationarity and non-stationarity in time series analysis," Physica D, vol. 99, (2-3), pp. 134-161, 1996.

[2] O Ochanda, " Time series and forecasting of monthly air temperature changes in Nairobi Kenya", M. S. thesis, UNIVERSITY OF NAIROBI, 2014.

[3] E. C. Nwogu, I. S. Iwueze and V. U. Nlebedim, "Some tests for seasonality in time series data," Journal of Modern Applied Statistical Methods, vol. 15, (2), pp. 24, 2016.

[4] I. S. Iwueze et al, "Uses of the Buys-Ballot table in time series analysis," Applied Mathematics, vol. 2, (05), pp. 633, 2011.

[5] C. Chatfield, " The analysis of time series: An introduction," Boca Raton, FL: Chapman & Hall/CRC,2004.

[6] S. Mukhopadhyay et al, "A comparative study between seasonal wind speed by Fourier and Wavelet analysis," ArXiv Preprint arXiv:1407.8476, 2014.

[7] B. Sule, O. Ajala, Investigating the Non-Stationarity of Historical Data on Rainfall at Ilorin, North Central Nigeria. The Journal of Scientific and Engineering Research, July, pp. 537-548, 2016.

[8] E. Afrifa-Yamoah, "Application of ARIMA models in forecasting monthly average surface temperature of Brong Ahafo Region of Ghana," *International Journal of Statistics and Applications,* vol. 5, *(5),* pp. 237-246, 2015.

[9] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *"Time series analysis: Forecasting and control*," Englewood Cliffs, NJ: Prentice Hall. 1994.

[10] N. Mohamed, M. H. Ahmad and Z. Ismail, "Short term load forecasting using double seasonal ARIMA model," in *Proceedings of the Regional Conference on Statistical Sciences, 2010.*

[11] H. Musbah, *M. Elhawary, " SARIMA Mo*del Forecasting of Short-Term Electrical Load Data Augmented by Fast Fourier Transform Seasonality Detection," IEEE Canadian Conference of Electrical and Computer Engineering, pp.766-770. 2019.