THE USE OF AUTOCORRELATION FUNCTION IN THE SEASONALITY ANALYSIS FOR FATIGUE STRAIN DATA

Z.M. Nopiah[1]
A. Lennie[2]
S. Abdullah[3]
M.Z. Nuawi[4]
A.Z. Nuryazmin[5]
M.N. Baharin[6]

ABSTRACT

The seasonal dependency or seasonality is a general component of the time series pattern that can be examined via correlograms, where the correlogram displays graphical and numerical information in an autocorrelation function. This paper discusses the use of an autocorrelation function in the seasonality analysis for the fatigue strain data to help identify the seasonal pattern. The objective of this study is to determine the capability of this time domain method in detecting the seasonality component in the fatigue time series. A set of case study data consisting of the non-stationary variable amplitude loading strain data that exhibits random behaviour was used. This random data was collected in the unit of micro-strain on a lower suspension arm of a mid-sedan car. The data was measured for 60 seconds at the sampling rate of 500 Hz, which gave 30,000 discrete data points. The collected data was then analysed in the form of an auto-correlogram shape and characteristics. As a result, an autocorrelation plot is weak in identifying the seasonal pattern, but the autocorrelation coefficient values are statistically significant and have shown a positive serial correlation. Thus, the finding of this characteristic is expected for a non-stationary signal.

[1] Fundamental Engineering Studies Unit Faculty of Engineering and Built Environment 43600 UKM Bangi, Selangor, Malaysia

[2] Fundamental Engineering Studies Unit Faculty of Engineering and Built Environment 43600 UKM Bangi, Selangor, Malaysia

[3] Department of Mechanical and Materials Engineering Faculty of Engineering and Built Environment 43600 UKM Bangi, Selangor, Malaysia

[4] Department of Mechanical and Materials Engineering Faculty of Engineering and Built Environment 43600 UKM Bangi, Selangor, Malaysia

[5] Fundamental Engineering Studies Unit Faculty of Engineering and Built Environment 43600 UKM Bangi, Selangor, Malaysia

[6] Fundamental Engineering Studies Unit Faculty of Engineering and Built Environment 43600 UKM Bangi, Selangor, Malaysia

INTRODUCTION

Fatigue failure is a process that involves crack initiation and propagation of a component under repeated loading. A signal is a series of numbers that come from measurement, typically obtained using some recording method as a function of time (Abdullah 2008). In the case of fatigue research, the data was measured based on a signal that consists of a measurement of cyclic loads, i.e. force, strain and stress against time (Meyer 1993).

Time series data can be described in a descriptive form as a set of data collected or arranged in a sequence of order over a successive equal increment of time (Lazim 2007). Over time, the behaviour of a time series can be described by characterising certain unique attributes, which can be identified and are generally grouped into four main component types (Lazim 2007). The identification of fatigue data behaviour is based on the existence of time series component which involves the identification of trend (T_t), cyclical (C_t), seasonal (S_t) and irregular (I_t) component in time period t. This method is called the classical decomposition of time series. Nopiah et al. (2009) examine the fatigue data behaviour based on this classical decomposition method and this study only reveals the trend, cyclical and irregular components which exist in the fatigue data signal.

This paper brings up the seasonality analysis in the variable amplitude fatigue data to determine a seasonal pattern if it exists in the fatigue strain data. In this study, the identification of seasonal variation is based on the autocorrelation function (ACF). There are three types of fatigue time series data used as a case study which present the strain signal of variable road surface. The objective of this study is to analyse the capability of this technique in detecting the pattern behaviour in terms of an autocorrelation plot in a strain-based time series on a lower arm suspension of a car.

LITERATURE BACKGROUND

Seasonal Component of Time Series

The seasonal component, also known as seasonal variation, refers to the characterization of regular fluctuations occurring within a specific period of time (Lazim 2007). The seasonal component is also referred as the seasonality of a time series. Seasonality (or periodicity) can usually be assessed from an autocorrelation plot, a seasonal subseries plot or a spectral plot. The following graphical techniques can be used to detect seasonality.

The run sequence plot is a recommended first step for analysing any time series. Although seasonality can sometimes be indicated with this plot, seasonality is shown more clearly by the seasonal subseries plot or the box plot. The seasonal subseries plot does an excellent job of

showing both the seasonal differences (between group patterns) and also the within-group patterns. The box plot shows the seasonal difference (between group patterns) quite well, but it does not show within the group patterns. However, for large data sets, the box plot is usually easier to read than the seasonal subseries plot. Both the seasonal subseries plot and the box plot assume that the seasonal periods are known.

Autocorrelation plots (Chatfield 1996) are a commonly-used tool for checking randomness in a data set. This randomness is ascertained by computing autocorrelations for data values at varying time lags. If random, such autocorrelations should be near zero for all time-lag separations. If non-random, then one or more of the autocorrelations will be significantly non-zero.

Autocorrelation Function (ACF)

The autocorrelation function (ACF) is an important diagnostic tool for analysing time series in the time domain (Rafiee & Tse 2009). The ACF is also very useful when examining stationarity and when selecting from various non-stationary models. In autocorrelation, lag is a time period separating the ordered data and is used to calculate the autocorrelation coefficients. The maximum number of lags is roughly $n/4$ for a series with less than 240 observations or ($\sqrt{n} + 45$) for a series with more than 240 observations, where n is the number of observations or for time series, the number of data points. When computed, the resulting number can range from +1 to -1. An autocorrelation of +1 represents a perfect positive correlation while a value of -1 represents a perfect negative correlation.

Let the time series of length N be x_t, $t = 1, ...,N$. The lagged scatterplot for lag k is a scatterplot of the last $N-k$ observations against the first $N-k$ observations (Chatfield 1996). Eq. 1 can be generalized to give the correlation between observations separated by k time steps, where $\bar{x} = \sum_{t=1}^{N} x_t$

is the overall mean. The quantity r_k is called the autocorrelation coefficient at lag k.

$$r_k = \frac{\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N-k}(x_t - \bar{x})^2} \qquad (1)$$

Autocorrelation plots, called the correlograms, present a better understanding of the evolution of a process through time using the probability of the relationship between data values separated by a specific number of time steps (lags). The correlogram plots autocorrelation coefficients on the vertical axis, and lag values on the horizontal axis. Correlogram summarizes characteristic features of the time series, viz. randomness, the rising or declining trend, oscillation, etc. Below are some keys for interpreting the correlogram (Chatfield 1996).

- *Random series—*If a time series is completely random, then for large N, $r_k = 0$ for all non-zero values of k.

- *Short term correlation*—Stationary time series exhibit short term correlations, characterized by a fairly large value of r_k followed by two or three more coefficients, while greater than zero tend to become successively smaller. Values of r_k tend to approximate zero for large k.

- *Alternating series*—If successive observations of a time series tend to alternate on different sides of the overall mean, the correlogram would also tend to oscillate.

- *Non-stationary time series*—If a time series has a trend, then the values of r_k would not decrease to zero, except for large values of k. However, it would be desirable to de-trend the series, otherwise its characteristics would be swamped by the trend.

- *Seasonal fluctuations*—If a time series is characterized by seasonal fluctuations, then the correlogram would also exhibit oscillations at the same frequency.

METHODOLOGY

A variable amplitude (VA) strain loading data was used in this study. It was collected from an automobile component during vehicle road testing. The fatigue data acquisition system, the SoMat eDaQ Data Acquisition, was used for the strain data measurements. The strain loading was measured using strain gauges of 5 mm size that was located at the maximum stress area. The signal was measured on the front left lower suspension arm of a car travelling over a pavéd road of 35–45 km/h, a highway route at 70–80 km/h and a campus road at 35–45 km/h. The road test data was collected for 60 seconds at a sampling rate of 500 Hz, which gave 30,000 discrete data points.

For the purpose of data comparison, the strain signal from the database of the Society of Automotive Engineers (SAE) profiles, i.e. the SAESUS (SAE standard suspension) was used. The SAESUS signal was collected from a suspension component of a car. It was assumed to be sampled at 204.8 Hz for 25,061 data points. It gave the total record length of the signal of 122.4 seconds. The analysis started with the extraction of data input from the data acquisition system of three different data profiles. Then, the seasonality analysis was carried out by plotting the correlogram of the autocorrelation function and the result recorded for interpretation and comparison purposes.

RESULT AND DISCUSSION

The captured non-stationary strain loading data was analysed in the time domain signal plots in Figure 1. Figure 1 represents the time history of the SAESUS data, pavéd road, highway and campus road. These fatigue signals have a variable amplitude (VA) pattern in the strain format. The SAESUS strain data gave the total record length of 122.4 seconds. It shows that the highest strain amplitude or strain range is 345 $\mu\varepsilon$.

Figure-1. Time history plots for: (a) SAESUS data, (b) pavéd, (c) highway, and (d) campus road
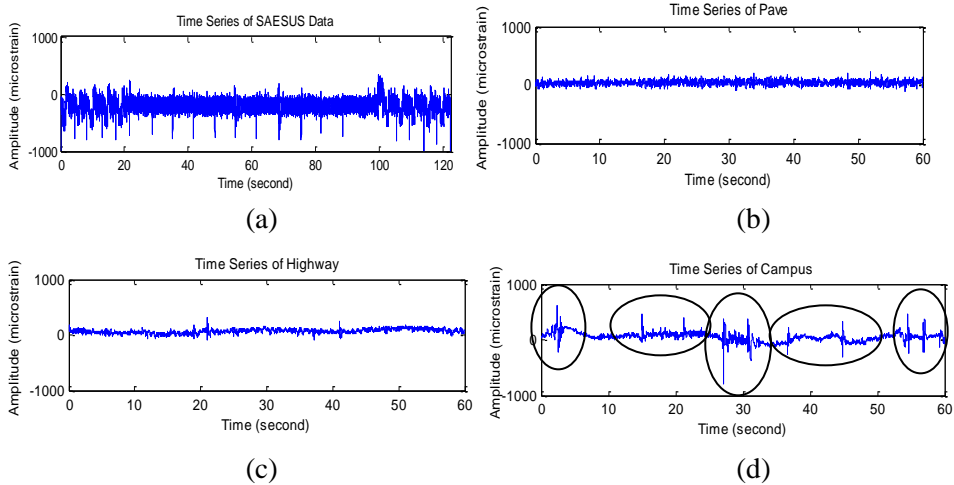
The study case data was measured on the suspension component of a car travelling on public road surface. The road load conditions were from a highway representing mostly consistent load features, a pavéd road representing noisy and consistent load features, and a campus road which represented load features that include turning and braking, driving on rough road surfaces and speed bumps. The time history plots show that the highest strain amplitude or strain range for the pavéd, highway and campus roads were 224 $\mu\varepsilon$, 321 $\mu\varepsilon$ and 619 $\mu\varepsilon$, respectively. Thus, it is found that the campus road has contributed the highest displacement to the lower suspension arm. Then, it is followed by the highway and pavéd roads. The potholes and bumps along the road also contribute to the spikiness of the fatigue data and this phenomenon is evident in Figure 2(d).

The seasonal patterns of the time series can be examined via the correlograms. Thus, in the data analysis, a correlogram serves as an image of the correlation statistics. Figure 2 illustrates the correlograms of each data set and it shows that these correlograms indicate the largest spike at lag 1. The SEASUS data correlogram in Figure 2(a) indicates that the shape is alternately positive and negative, and this plot can be considered as a seasonal fluctuation. The pavéd correlogram in 2(b) has indicated that the shape is alternating positively and negatively correlation, decaying to zero and this plot can be considered as a short term correlation. Meanwhile, the correlograms for the highway and campus road in Figures 2(c) and 2(d) demonstrate strong and positive autocorrelations and the autocorrelation decays very slowly, and this is a non-stationary time series. The values of the autocorrelation coefficient at lag 1, r_1 are 0.969, 0.972, 0.992 and 0.996 for SAESUS, pavéd, highway and campus, respectively. The r_1 values are statistically significant and they show a positive serial correlation. By observing these correlograms, it can be assumed that the road surface difference can affect the strain signal produced and the stationary behaviour.

Figure-2. The correlograms for: (a) SAESUS data, (b) pavéd, (c) highway, and (d) campus road

(a)

(b)

(c)

(d)
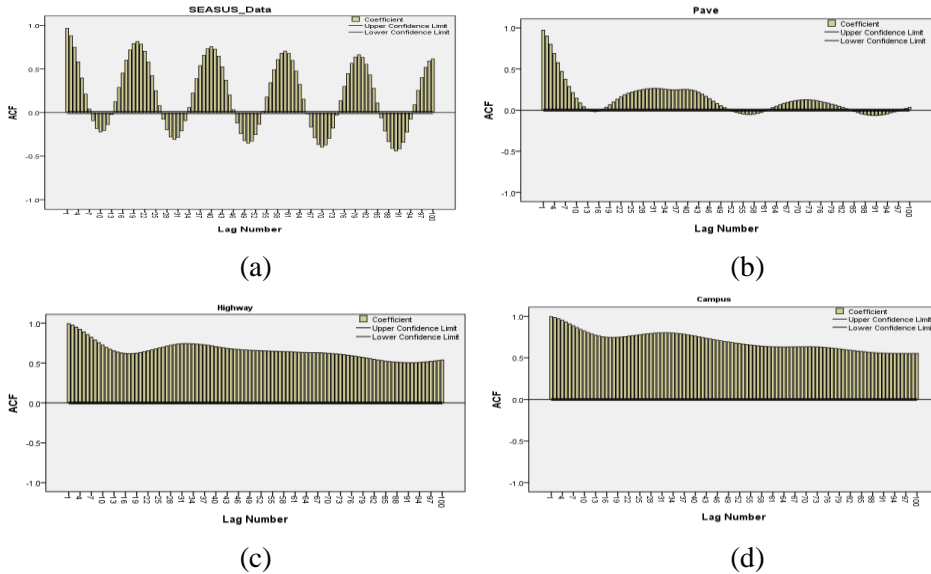
CONCLUSION

This work elaborates on the seasonal analysis on variable amplitude (VA) loading strains data by using the autocorrelation function (ACF). This method is used to observe a seasonal pattern if it exists in the fatigue time series data. A set of case study data consisting of non-stationary VA loading strains data that exhibit a random behaviour was used and then the case study data was then compared with the SAE standard suspension, i.e SEASUS data. The time history plot has displayed the fatigue data containing both high and low amplitudes and these characteristics signify a nonstationary behaviour. The non-stationary behaviour of the fatigue data can be supported with the correlogram, when a significant larger spike followed by smaller spikes at lag 1 and the pattern of correlogram slowly decays to zero. The seasonality analysis using the ACF is an initial step to analyse the non-stationary behaviour of a signal. This method is weak in identifying the high and low amplitude events, and it can only observe the non-stationary behaviour via the correlogram pattern. This paper can propose a time-frequency domain method that can be used for further study to make better identification of the non-stationary behaviour.

ACKNOWLEDGEMENT

REFERENCES

Abdullah, S., Ibrahim, M.D., Nopiah, Z.M., Zaharim, A. (2008) Analysis of a variable amplitude fatigue loading based on the quality statistical approach, Journal of Applied Sciences Vol.**8**, pp.1590-1593.

Chatfield, C. (1996) The analysis of time series. Fifth Eds., Chapman & Hall.

Lazim, M.A. (2007) Introductory Business Forecasting, A Practical Approach, Second Eds., University Publication Centre (UPENA).

Meyer, Y. 1993. Wavelets: Algorithm and Applications. SIAM, Philadelphia.

Nopiah, Z.M., Baharin, M.N., Abdullah, S., Khairir, M.I. (2009) Time Series Behaviour of Lower Arm Suspension Fatigue Data using Classical Decomposition Method, 2009 International Conference on Signal Processing Systems, pp. 984 – 988, 15-17 May 2009, Singapore.

Rafiee, J. & Tse, P.W. (2009) Use of autocorrelation of wavelet coefficients for fault diagnosis, Mechanical Systems and Signal Processing Vol.**23**, pp.1554–1572.