SMAR 2019

Application of Time Series Methods on Long-Term Structural Monitoring Data for Fatigue Analysis

Morteza AHMADIVALA [1], Bartłomiej SAWICKI [2], Eugen BRÜHWILER [2], Thierry YALAMAS [1], Nicolas GAYTON [3], Cécile MATTRAND [3], André ORCESI [4]

[1]PHIMECA Engineering, Cournon d'Auvergne, France
[2]Laboratory of Maintenance and Safety of Structures: École polytechnique fédérale de Lausanne, Lausanne, Switzerland
[3]Université Clermont Auvergne, CNRS, SIGMA Clermont, Institut Pascal, Clermont-Ferrand, France
[4]IFSTTAR: L'Institut français des sciences et technologies des transports, de l'aménagement et des réseaux, Marne-la-Vallée, France

Contact e-mail : ahmadivala@phimeca.com

ABSTRACT: Structural health monitoring (SHM) can be employed to reduce uncertainties in different aspects of structural analysis such as: load modeling, crack development, corrosion rates, etc. Fatigue is one of the main degradation processes of structures that causes failure before the end of their design life. Fatigue loading is among those variables that have a great influence on uncertainty in fatigue damage assessment.

Conventional load models such as Rain-flow counting and Markov chains work under stationarity assumption, and they are unable to deal with the seasonality effect in fatigue loading. Time series methods, such as ARIMA (Auto-Regressive Integrated Moving Average), are able to deal with this effect in the data; hence, they can be helpful for fatigue load modelling. The goal of this study is to implement seasonal ARIMA to prepare a load model for long-term fatigue loading that can capture more details of the loading scenario regarding the seasonal effects in traffic loading.

1 INTRUDUCTION

Structural Health Monitoring (SHM) plays an important role for the maintenance and inspection of existing structures. It helps in understanding the structure's condition for a better allocation of maintenance funding by reducing the uncertainties in degradation processes, like fatigue and corrosion among others (Long et al. (2018)). Motorized traffic loading introduces large uncertainties due to its stochastic nature. An appropriate load model is necessary for realistic structural analysis. Therefore, the main objective of this study is to provide the fatigue load model on the basis of long-term monitoring data by means of ARIMA (Auto-regressive integrated moving average).

SHM can be utilized to provide different types of information about the structure, and the monitoring duration is crucial. Monitoring of a longer duration can provide more precise information on the structure thanks to capturing the extreme events, seasonal effects in structural loading, etc. But the gaps due to limited lifetime of gauges, power outages and computer errors are difficult to avoid. On the other hand, processing and utilizing the data is another issue as its size can be enormous. The proposed method can help in dealing with these problems.

Several approaches can be identified in the literature to prepare a load model for structural fatigue analysis among which: cycle counting methods, like rain-flow counting; and random process methods, like Markov chain method, are most utilized (Khosrovaneh & Dowling (1990)). It should be noted that the load amplitudes are discretized into different levels in both approaches (Krenk & Gluver (1989)). Frequency domain-based and time domain-based methods can be used to provide the load model as a random process with a continuous state space. Application of the frequency domain-based methods for fatigue load modeling is complicated since all operations against fatigue, such as risk assessment, maintenance and inspection planning are performed in the time domain. Hence, time series methods that are defined in time domain can provide a more suitable load model. The prediction accuracy of ARMA (Auto Regressive Moving Average) as a time series method was compared with rain-flow counting and Markov chain methods (Ling et al. (2011)). All the methods have acceptable performance based on Bayesian hypothesis testing, however ARMA performs better than the other methods.

It is worthy to mention that all previous methods are applicable on stationary data. However, in reality the fatigue loading might have a non-stationary behavior, especially because of seasonality in traffic and environmental loadings. Seasonality refers to a seasonal pattern that can occur because of seasonal factors, such as the day of the week or time of the year. For instance, fatigue loading can experience a seasonality because of weekends when, the heavy traffic is reduced (Treacy (2014)). Among different time series methods, seasonal ARIMA is a strong tool that can be used for fatigue load modeling, and it is capable of dealing with seasonality. The goal of this research is to employ seasonal ARIMA to prepare a better load model for fatigue analysis being able to capture the seasonality effect. However, applying the seasonal ARIMA is not very straightforward on long-term loading data due to many observations within the seasonal window, understood as a period when the seasonal pattern occurs.

The reminder of this article is organized as following: in Section 2 time series methods are reviewed; Section 3 is related to the long-term monitoring data used in this study; the proposed approach for load modelling using long-term monitoring data is introduced in Section 4 and a short conclusion is provided in Section 5.

2 TIME SERIES METHODS: ARIMA

2.1 ARIMA

ARIMA is a time series method that combines Auto-Regressive (AR) and Moving Average (MA) models integrated with differencing. AR predicts the current value of a desired variable by a linear combination of the past values of that variable (Hyndman & Athanasopoulos, (2018)). The following equation shows an AR model of order *p*

$$y_t = c + \varphi_1 y_{t-1} + \cdots + \varphi_p y_{t-p} + \varepsilon_t \tag{1}$$

where: y_t is the predicted value of the desired variable, c is a constant, ε_t is white noise, and $\varphi_i: i = 1, \dots, p$ are the coefficients of the AR model of order p.

The MA model uses the errors ε of the past values regarding to the mean value of the data μ in a regression-like model to predict the current value for the desired variable according to equation 2.

$$y_t = \mu + w_1 \varepsilon_{t-1} - \cdots - w_q \varepsilon_{t-q} \tag{2}$$

$w_i: i = 1, \dots, q$ are the coefficients of the MA model of order q.

An ARIMA model of order (p,d,q) (equation 3) is constructed by combining an AR of order p with MA of order q, and d is the degree of differencing.

$$y_t^{(d)} = c + \varphi_1 y_{t-1}^{(d)} + \cdots + \varphi_p y_{t-p}^{(d)} + w_1 \varepsilon_{t-1} + \cdots + w_q \varepsilon_{t-q} + \varepsilon_t \tag{3}$$

2.2 Seasonal ARIMA

Seasonal ARIMA can be constructed by adding the extra seasonal term into the non-seasonal ARIMA and it can be formulated as follows:

$$seasonalARIMA\text{:} \quad \underbrace{(p, d, q)}_{non-seasonal\ part} \quad + \quad \underbrace{(P, D, Q)_m}_{seasonal\ part} \tag{4}$$

P, D, and Q represent the parameters for the ARIMA model of seasonal behavior with m observations for each seasonal window. By removing the seasonal behavior, a non-seasonal ARIMA model with parameters p, d and q can be fitted to the residuals. In fact, the seasonal effect is removed from the data for sake of stationarizing, and the seasonal pattern can be modeled with another ARIMA (Hyndman & Athanasopoulos, (2018)).

2.3 Identification of the best ARIMA model

The first step to construct the ARIMA model is to find the orders p, q and the differencing degree d. The main purpose of application of the differencing is to stationarize the data. Hence, the best value for d is such that gives a stationary residual after differencing the data d-many times. Akaike's Information Criterion (AIC) is used to identify the orders p and q of an ARIMA model. AIC criteria can be formulated as:

$$AIC = -2log(L) + 2(p + q + k + 1) \tag{5}$$

Where L is the likelihood of the time series and k=0 if c=0, otherwise k=1 (see equation 3).

2.4 Parameter estimation using maximum likelihood

After identifying the order of an ARIMA model, associated parameters $(c; \varphi_i; w_j; i = 1, \dots, p; j = 1, \dots, q)$ need to be approximated. One of the most common methods to estimate the parameters is Maximum Likelihood Estimation (MLE). MLE tries to approximate the parameters in a way such that the probability to obtain the observed data is maximized. MLE for ARIMA models is similar to least squares estimates that tries to minimize the square error between the observed and approximated data.

2.5 Model validation

One common way to test the model is to divide the available data into two parts: training and test data. Training data is used to fit the ARIMA and to estimate the model parameters. After defining the ARIMA model, it will be validated on test data to evaluate the accuracy of the model. More details about time series methods can be found in (Hyndman & Athanasopoulos, (2018)).

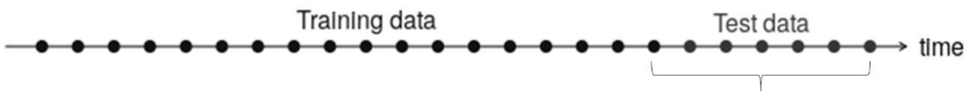

Figure 1. Scheme of training and test data concept.

3 LONG-TERM MONITORING DATA

In this study, the long-term monitoring data of Chillon viaducts was used. The Chillon viaducts are two parallel post-tensioned concrete structures in the Geneva lake region, Switzerland. The monitoring was commenced in 2016, after upgrading the structure with UHPFRC (Ultra High-Performance Fiber Reinforced Cementitious composite) (Brühwiler et al. (2015), Martín-Sanz et al. (2018)). In this paper, the data coming from the strain gauges installed on the transversal rebars was used. The data is available for almost two years with some gaps. During the monitoring campaign, the signals from the strain gauges have been recorded with the frequency of 50, 100, or 200 Hz, depending on period.

Figure 2A shows an example of the raw data from one full day. The wave in data comes from the difference in temperature between day and night. Figure 2B shows the data after removing this thermal effect (Treacy (2014)). Additionally, only strain-reversal points of cycles bigger than 1 micro-strain were kept to reduce the size of data.

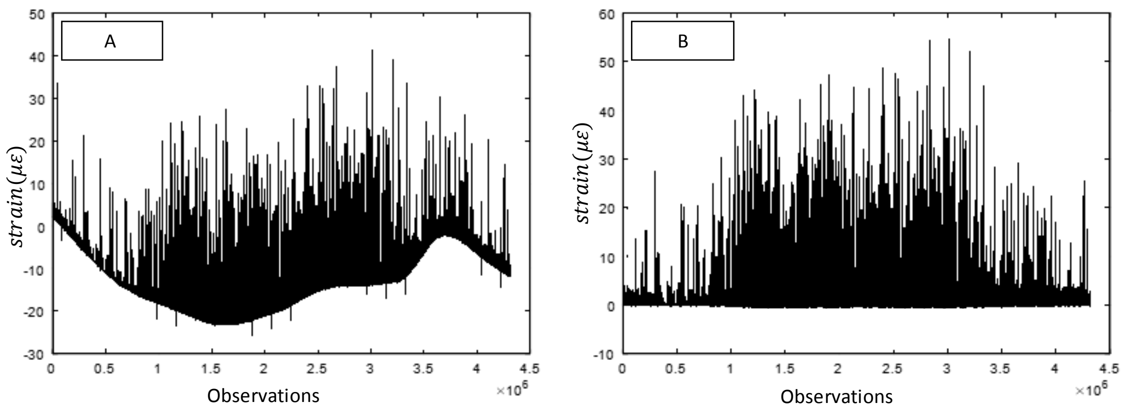

Figure 2. A: raw data from the strain gauge from midnight to midnight, B: processed data from the strain gauge from same period.

In the next section seasonal ARIMA is going to be used to prepare the load model utilizing the data from Autumn 2016. The reason of choosing this period was small amount of the monitoring interruptions. Since there is a linear relationship between the strain and stress values based on the Hook’s law, the ARIMA model has been directly applied on the observed strains.

4 APPLYING ARIMA TO LONG-TERM STRAIN DATA

4.1 Challenge

As it was mentioned before employing the ARIMA method on long-term monitoring data is not very straightforward. The difficulty is that the number of observations within a seasonal window is very large. If we consider that the seasonality window for traffic loading is one week (since there is a big difference between the loading experienced during the weekdays and weekends), the number of load cycles recorded in this period would be considerable (more than 10’000 cycles). This makes the process of applying seasonal ARIMA on the long-term load monitoring very difficult as the allowable observations within a seasonal window in conventional applications is around 350. However, in practice, for seasonal window larger than 200 observations available models in R and Python run out of memory. Hence, in the following part the data is treated differently to apply the seasonal ARIMA.

4.2 Proposed solution

To tackle the issue of a long seasonality window, the long-term monitoring data needs to be treated differently. For this reason, instead of considering each maximum or minimum value in the load data as a sample for time series analysis, the monitoring period was divided into shorter periods, such as monthly, weekly or daily. The aim is to define a probability distribution for the load cycles within each time interval. The parameters of the distributions will be used to replace the observations in the time series analysis. Hence, a few values (like mean value and standard deviation) can be employed to represent the big amount of observations in each period. Finally, time series analysis can be performed on the parameters of the distributions.

This transformation is well represented in Figures 3 and 4 where the former shows the strain values recorded for autumn 2016 and the latter shows the mean values for days and nights during autumn. The time interval used for this study is 12 hours, which divide the monitoring period into days and nights. This discretization strategy is used as the traffic flow during the day and night is different, and it is a reasonable period where there is no seasonality effect in the observations. Seasonal ARIMA can be used to prepare a load model for the monitoring data after this transformation.

4.3 Seasonal ARIMA results

After the data transformation according to the previous section, seasonal ARIMA was applied on the mean values of the monitoring data for each 12 hours. As explained before, in seasonal ARIMA, the seasonal part of the data is removed first. It can be done by seasonal differencing as in the following equation, where m is the seasonality window. For our data it is equal to 14 (nights and days in a week).

$$y'_t = y_t - y_{t-m} \tag{7}$$

Two ARIMA models are fitted to the seasonal data and the residuals (as in equation 4). Figure 5 illustrates the Q-Q plot and the correlogram on the residuals, showing that the seasonality is well removed and the residuals are fairly stationarized. KPSS (Kwiatkowski-Philips-Schmidt-Shin) stationarity test has been used to check the stationarity of the residuals. P-value of 0.79 confirms the good level of stationarity while the null hypothesis in KPSS is that the data is stationary. Table 1 shows the best ARIMA order that has the minimum AIC among other orders, and Table 2 show the calculated parameters for the model using maximum likelihood estimation.

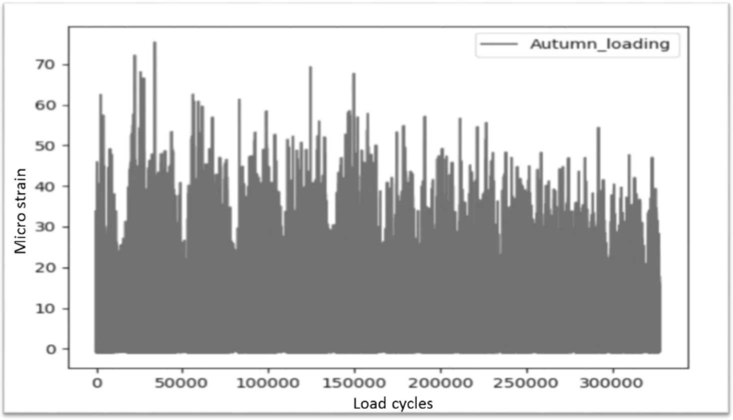

Figure 3. Strain values recorded for autumn 2016.

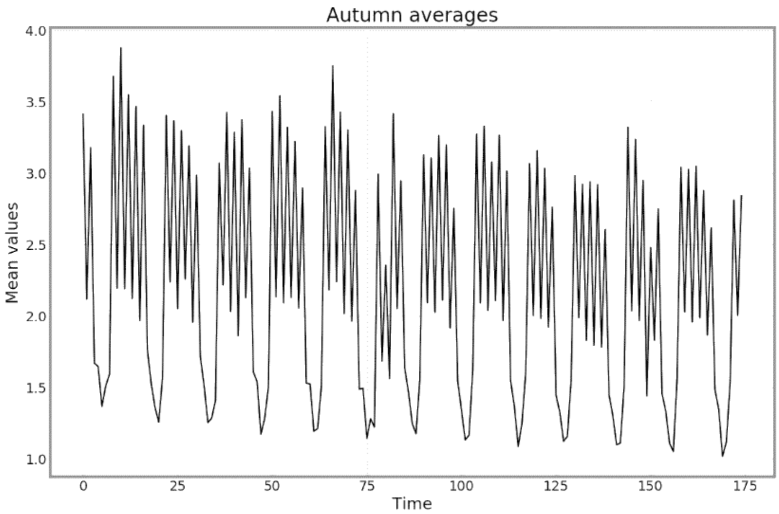

Figure 4. Mean values for days and nights for autumn 2016.

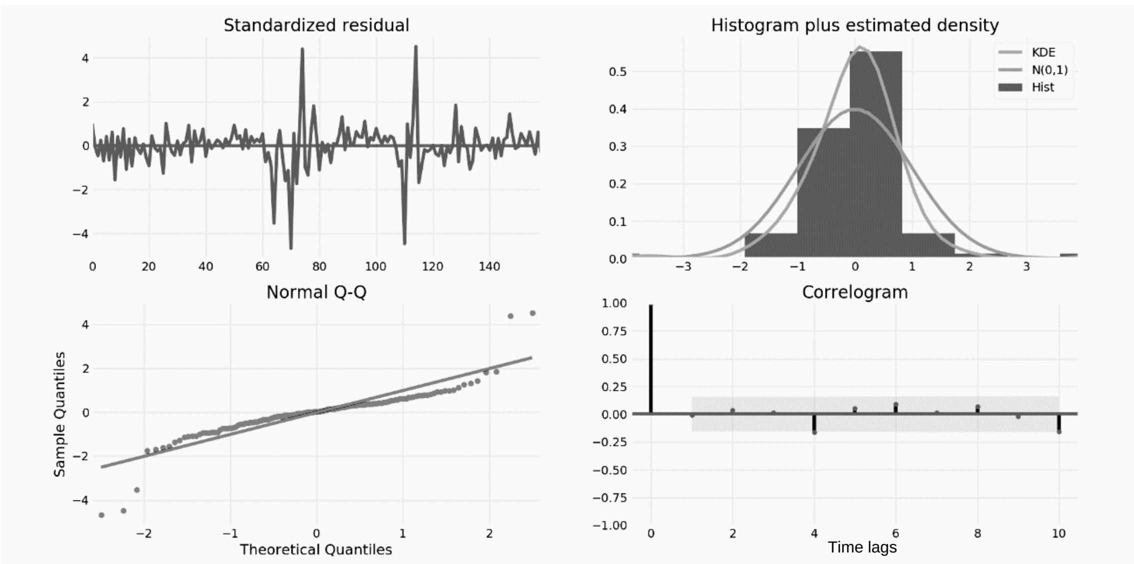

Figure 5. Q-Q plot and autocorrelation on residual data.

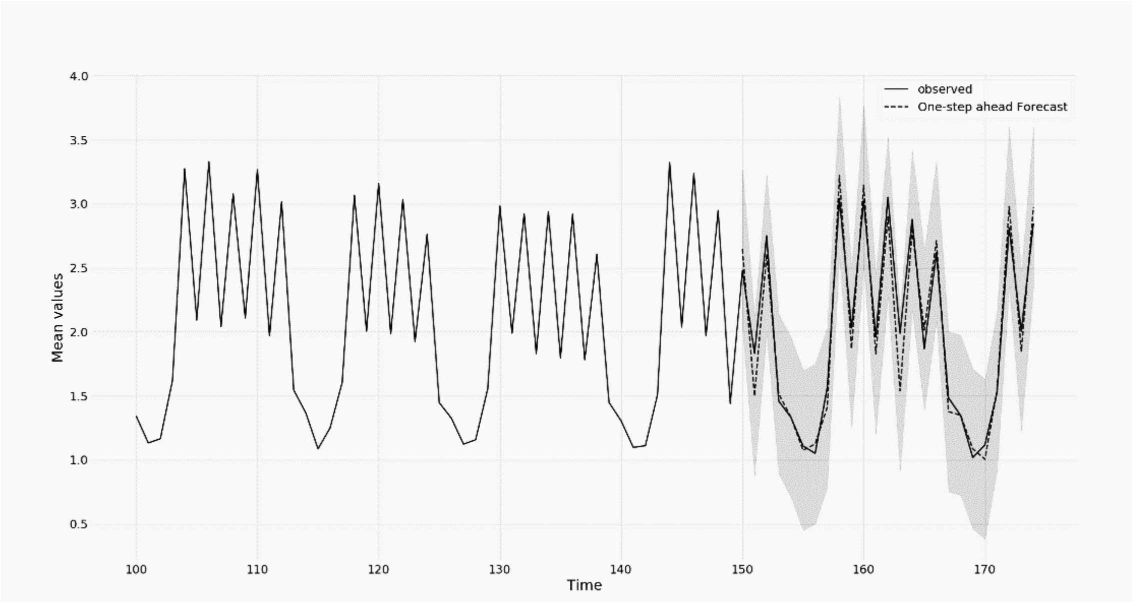

Figure 6. Model validation.

Table 1. Best order for seasonal ARIMA using the minimum AIC

p	d	q	P	D	Q	m	AIC
2	0	1	1	0	1	14	86.2

Table 2. Parameters for best seasonal ARIMA

φ_1	φ_2	w_1	φ_{1s}	w_{1s}
-0.38	0.62	0.99	0.96	-0.11

The last step is related to the model validation. As illustrated in Figure 6, some part of the data is used as the training data and the rest as test data. The seasonal ARIMA prepared for this data and then this model is used to provide predictions for the test data. Comparison between the test data and the model predictions is provided by means of the 90% confidence interval for one step ahead forecasting (the shaded area). The one step ahead forecasting uses the previous observations to predict y_{t+1}. It can be seen that all the predictions are close enough to the test data and all of them are within the desired confidence interval.

The suggested seasonal ARIMA model in this section can provide a useful tool to reproduce the loading scenario on different structures for further applications such as structural fatigue life assessment. What makes this model different from previous models is that it can capture the seasonality effect in the loading and is easy to employ on the long-term monitoring data. On the other hand, this model can be used to deal with the missing monitoring data and to predict the future loading, since the seasonality effect is implemented inside this model.

5 CONCLUSION AND PERSPECTIVES

Application of seasonal ARIMA on long-term monitoring data of traffic load effects is studied in this paper. It can help in understanding the seasonal behavior of the structure under motorized traffic. The difficulty related to the seasonality window on long-term loading data is its big size, which makes it very difficult to apply the seasonal ARIMA. The data transformation has been introduced in this study to tackle this difficulty. The new approach has been implemented on the long-term monitoring data on Chillon viaduct. Part of monitoring data was used for the model calibration, while the remaining for verification. A very good agreement between foreseen and measured values was obtained.

The proposed algorithm for load model preparation can be used not only for the fatigue life assessment, but also to deal with the missing monitoring data and to predict the future loading scenarios. To achieve this, the metamodels to transform the distribution time series into the load cycles should be further developed.

6 REFERENCES

Brühwiler, E., Bastien-Masse, M., Mühlberg, H., Houriet, B., Fleury, B., Cuennet, S. Maurer, M., 2015. Strengthening the Chillon viaducts deck slabs with reinforced UHPFRC. In *IABSE Symposium Report* (Vol. 105, pp. 1–8)

Hyndman, R.J., & Athanasopoulos, G. 2018. Forecasting: principles and practice, 2nd edition, Melbourne, Australia. OTexts.com

Khosrovaneh, A. K., & Dowling, N. E. 1990. Fatigue loading history reconstruction based on the rainflow technique. *International Journal of Fatigue*, 12(2), 99–106

Krenk, S., & Gluver, H. 1989. A Markov matrix for fatigue load simulation and rainflow range evaluation. *Structural Safety*, 6(2), 247–258

Ling, Y., Shantz, C., Mahadevan, S., & Sankararaman, S. 2011. Stochastic prediction of fatigue loading using real-time monitoring data. *International Journal of Fatigue*, 33(7), 868–879.

Long, L., Thöns, S., & Döhler, M. 2018. The effects of SHM system parameters on the value of damage detection information. *EWSHM; European workshop on structural health monitoring*, Manchester, United Kngdom,

Martín-Sanz, H., Avendano Valencia, L. D., Dertimanis, V., Chatzi, E., & Brühwiler, E. 2018. Monitoring of Chillon viaduct after Strengthening with UHPFRC. Presented at *the 9th International Conference on Bridge Maintenance, Safety and Management (IABMAS 2018)*, Melbourne, Australia

Treacy, M. A., 2014. The use of monitored data in the verification of structural and fatigue safety of existing post-tensioned concrete highway bridges. Doctoral Thesis, École polytechnique fédérale de Lausanne, Lausanne, Switzerland

7 AKNOWLEDGMENT

This project has received funding from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 676139.
Many thanks to Bahman Madadi, PhD researcher at TU Delft, for his helps towards this study.