

Rail Degradation Modelling Using ARMAX: A Case Study Applied to Melbourne Tram System

M. Karimpour, N. Elkhoury, L. Hitihamillage, S. Moridpour, R. Hesami

Abstract—There is a necessity among rail transportation authorities for a superior understanding of the rail track degradation overtime and the factors influencing rail degradation. They need an accurate technique to identify the time when rail tracks fail or need maintenance. In turn, this will help to increase the level of safety and comfort of the passengers and the vehicles as well as improve the cost effectiveness of maintenance activities. An accurate model can play a key role in prediction of the long-term behaviour of railroad tracks. An accurate model can decrease the cost of maintenance. In this research, the rail track degradation is predicted using an autoregressive moving average with exogenous input (ARMAX). An ARMAX has been implemented on Melbourne tram data to estimate the values for the tram track degradation. Gauge values and rail usage in Million Gross Tone (MGT) are the main parameters used in the model. The developed model can accurately predict the future status of the tram tracks.

Keywords—ARMAX, Dynamic systems, MGT, Prediction, Rail degradation.

I. INTRODUCTION

IN the past few decades, the expansion of transport networks has been the main concern of transport organizations. Since the infrastructure of railways is growing fast, maintenance planning became their main focus to ensure the operation of rail systems at best and safe practice standards.

Currently, rail transport organizations heavily rely on the experience and expertise of the maintenance teams to identify and locate the segments on rail tracks, which need to be repaired or replaced. However, manual inspections are costly and time consuming. Therefore, the focus has shifted onto intelligently managing the existing rail tracks, and efforts have been made to design a structured framework for maintenance planning and help reducing the unnecessary maintenance costs. Previous studies have mostly focused on modelling and prediction of rail track degradation for heavy rail system (train) rather than light rail system (trams) [1]-[4].

This particular study will focus on tram network of Melbourne, Australia, which is the largest metropolitan tram network in the world and covers 250 km of rail tracks [5]. The

M. Karimpour is a research assistant at School of Civil, Environmental and Chemical Engineering, RMIT University, Melbourne, Australia (e-mail: mostafa.karimpour1@gmail.com).

N. Elkhoury and L. Hitihamillage are masters' students at School of Civil, Environmental and Chemical Engineering, RMIT University, Melbourne, Australia (e-mail: s3272253@student.rmit.edu.au, s3375355@student.rmit.edu.au).

S. Moridpour is a senior lecturer at School of Civil, Environmental and Chemical Engineering, RMIT University, Melbourne, 3001, Australia (phone: +61-3-9925-2407, fax: +61-3-9639-0138, e-mail: sara.moridpour@rmit.edu.au).

R. Hesami is an advisor, Asset planning and visualization at Yarra Trams, GPO Box 5231 Melbourne, Australia (e-mail: rihannah.hesami@yarratrams.com.au).

data for the Melbourne tram network have been collected from 2010 to 2015 through inspection on-sight and stocked in a non-digitized way for a long time. Different parameters contribute to the degradation of rail tracks. The gauge and MGT are important parameters for tram tracks degradation. In this context, an autoregressive moving average with exogenous input (ARMAX) is developed for curves sections using the gauge and MGT values to predict the degradation of tram tracks over the next years. Hence, this model allows decision makers to achieve an optimization in maintenance activities and to save time and costs.

The structure of the paper is as follows. Section II presents a relevant literature review on previous degradation models. Section III proposes an ARMAX model covering the noise of the data with a very low Mean Absolute Error (MAE). Section IV concludes by suggesting directions for future research.

II. LITERATURE REVIEW

A number of previous researches presented different models of railway degradation prediction [1]-[4]. Different parameters contributed to the tram tracks degradation modelling such as asset condition (i.e. sleepers, ballast, fastening) [6], age of rails, axle load [7], [8], speed [8], MGT [7], track curvature [8], [9] and rail lubrication [10]. According to the previous studies, rail degradation was predicted using different model types. One of these model types is the statistical models, which are based on parameters of the rail structure and the affecting factors of the tram tracks such as traffic and track components. Andrade and Teixeira proposed Hierarchical Bayesian Models (HBM), which are flexible statistical models that provide a prediction of the railway degradation [11]. The study was considering two main quality parameters, which are longitudinal level defects and horizontal alignment defects, in relation to the degradation of rail track geometry. The structure of this model adopts the quality parameters as random variables that can be uncertainly calculated by a prior distribution. This study concluded that horizontal alignment defects are less predictable than longitudinal level defects. HBMs were also developed by other research studies [12]-[14]. However, these models rely on other statistical models such as Markov models, especially in the case of high numerical data [15], [16].

Another degradation model type is stochastic models, which are also statistical models. They aim to understand the distribution of time to degradation events and predict their performance. A stochastic model was developed to predict the degradation of the Portuguese railway Northern Line [17]. This study showed a high accuracy of the model developed for

the rail degradation over a period of time. However, the application of this model may require more understanding and explanation of the process. Also, there is no evidence to clearly validate the claim of the degradation pattern and its distribution.

Artificial Neural Networks (ANNs) are another type of modelling that has been used to predict the degradation of railways [18], [19]. ANNs refer to the knowledge of biological neural networks. They estimate functions depending on large numbers of inputs and unknowns. Reference [20] used different parameters to predict the rail degradation using an ANNs model which include: the combined track record index (CTR), traffic volume (e.g. light and heavy), speed, geographic location (e.g. plain, hilly, and mountainous), curves radius and gradient. The study compared the model predictions to the observed data of one of the sets. Consequently, this comparison showed that the following year CTR indices were at the same level as the CTR indices of the previous year or slightly lower than that. Furthermore, another study was proposed presenting an ANNs model to predict the degradation of tram tracks using maintenance data in Melbourne [21]. The data were collected and divided into three categories including inspection data, load data and repair data. Inspection data were collected for Melbourne tram network, from 2009 to 2013, covering different types of segments of four routes such as straights, curves, H-crossings and crossovers. Out of these segments, curves were the focal point since they have a higher failure rate than the other segments [22], [23]. This study showed that the load data without passengers in MGT, the gauge values and the rail profile were the most affecting factors of the ANNs degradation model. The application of the ANNs model was highly accurate.

Based on this review, a comparison of different degradation models was conducted. The majority of previous degradation models were general and oriented towards different parameters and variables affecting the degradation of rail tracks which is not available in Melbourne track data. As a result, in this paper, we suggest an ARMAX model focused on the gauge and MGT to predict the degradation of tram tracks in Melbourne over the next years. This will help predict the future maintenance procedures needed for the tram tracks resulting in fewer expenses, less effort and time saving.

III. METHODOLOGY

The rail geometric degradation is usually quantified with many different defects. One of the most important defects is gauge deviation defect and the MGT factor. In this paper, the main focus is the prediction of gauge defect in the future. The entire data used for this research consist of gauge and MGT values from 2010 to 2015, respectively. In Fig. 1, Gauge is plotted in respect to MGT. When Fig. 1 is considered, it is obvious that there is a meaningful relationship between gauge and MGT. Therefore, the gauge model $gauge(t+1)$ is considered to be a function of $gauge(t-i)$ for $i=0$ to n and $MGT(t-i)$ for $i=0$ to n the function could be

described as (1):

$$Gauge(t+1) = f(gauge(t-i), MGT(t-i)) \text{ for } i=0,1,\dots,n \quad (1)$$

To model the system, ARMAX approach has been utilized in this research. Statisticians use the time series terminology to explain applications that have sequenced successive equally spaced points in time. Some examples of time series modelling are the world oil price, daily price of power market, weather forecast and railway deterioration prediction [24]-[30]. Railway deterioration is affected by time and different parameters (e.g. the longitudinal levelling defects, the horizontal alignment defects, the cant defects, the gauge deviations and the track twist). However, these variables are mainly not provided in the datasets. Therefore, the analysis of this study is mostly focused on the gauge defect and MGT variables. The gauge is modelled as the most important factor of tram tracks degradation.

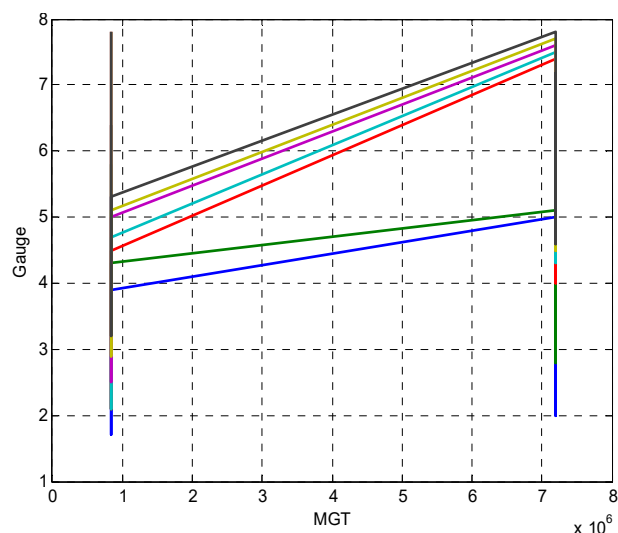


Fig. 1 The relationship between gauge values and MGT

The Autoregressive (AR) time series model is a very common type of system representation with few linear parameters [30]. In the AR model, output of system is derived in an autoregressive manner to previous value of outputs by filtering the white noise $v(k)$ as shown in (2):

$$y(k) = \frac{1}{D(q)} v(k) \quad (2)$$

$y(k)$ is the output in time k , and $v(k)$ is the white noise. It is clear that time series models are not accurate enough without considering the input so autoregressive with exogenous input (ARX) model is the extended form of the AR model which can be written as (3):

$$y(k) = \frac{B(q)}{A(q)} u(k) + \frac{1}{D(q)} v(k) \quad (3)$$

As a result of high noise in the gauge values, different dynamic system models have been evaluated on the data and ARMAX model was best fitted with the lowest MAE. Therefore, this method has been selected for system modelling. ARMAX model is in fact the extended noise model of the ARX with more flexibility. The ARMAX model is one of the most useful models in linear dynamic system modelling, although the model is nonlinear in parameters. ARMAX can be described as shown in (4):

$$A(q)y(k) = B(q)u(k) + C(q)v(k) \quad (4)$$

The predictor of ARMAX can be written as (5):

$$\hat{y}(k|k-1) = \frac{B(q)}{C(q)}u(k) + \left(1 - \frac{A(q)}{C(q)}\right)y(k) \quad (5)$$

The predictor is stable if $C(q)$ is stable. The prediction error of ARMAX model can be written as:

$$\epsilon(k) = \frac{A(q)}{C(q)}y(k) - \frac{B(q)}{C(q)}u(k) \quad (6)$$

The estimation of the ARMAX model can be done through the following procedure. First, an ARX model estimation for the data should be calculated as shown in (7) with respect to the data.

$$\hat{\phi}_{ARX} = (X^T X)^{-1} X^T y \quad (7)$$

In the second part, the ARMAX model parameters should be calculated with a nonlinear procedure. By using nonlinear least square methods, the model parameters can be identified. For the nonlinear least square models the computation of the gradients is necessary.

As the squared error is $e^2(k) = (y(k) - \hat{y}(k))^2$, so $\frac{\partial e^2}{\partial \phi} = -2\epsilon(k) \frac{\partial \hat{y}}{\partial \phi}$. Thus, the gradient of the estimated model must be calculated.

By multiplying both sides of (5) by $C(q)$, the equation could be rewritten as:

$$C(q)\hat{y}(k|k-1) = B(q)u(k) + (C(q) - A(q))y(k) \quad (8)$$

The differentiation $\frac{\partial e^2}{\partial \phi}$ yields to the differentiation (8) with respect to a_i, b_i, c_i .

$$C(q) \frac{\partial \hat{y}(k|k-1)}{\partial a_i} = -y(k-i) \quad (9)$$

Therefore,

$$\frac{\partial \hat{y}(k|k-1)}{\partial a_i} = -\frac{1}{C(q)}y(k-i) \quad (10)$$

Equation (9) should be calculated with respect to b_i, c_i which yields to:

$$C(q) \frac{\partial \hat{y}(k|k-1)}{\partial b_i} = u(k-i) \quad (11)$$

and,

$$\frac{\partial \hat{y}(k|k-1)}{\partial c_i} = \frac{1}{C(q)}y(k-i) - \hat{y}(k-i|k-i-1) \quad (12)$$

Therefore, the gradient can be calculated by the above equations. Various experiences have reported that the above equation convergence to a global optimal parameters [29].

In following section, ARMAX approach is implemented on the Yarra tram datasets on curve sections. The model has been trained on each sample and tested on the rest of the data. Fig. 2 shows the MAE of model on the data when sample x has been used as the training data; the MAE is the error calculated on all the data excluding the trained data.

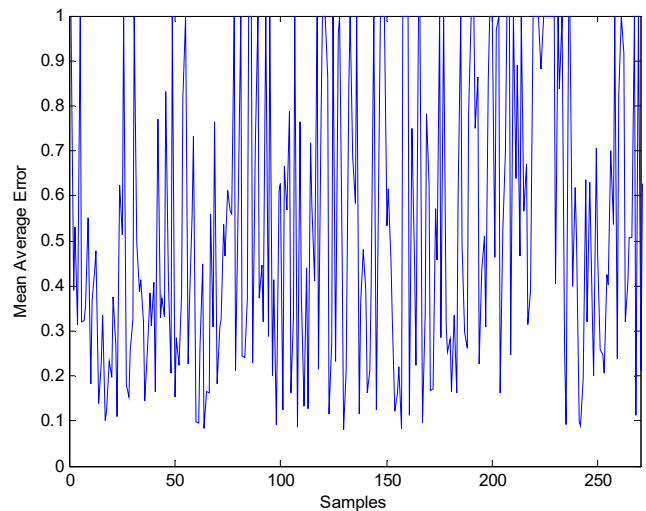


Fig. 2 MAE on test data when sample x has been used as the training data

To find the best sample that can be used in modelling the system, R-square criteria need to be calculated. Fig. 3 demonstrates the R-square value for the test data if sample x is used to model the system.

Fig. 3 shows the R-square value of the model is different when it is trained using different samples. In some samples, the R-square is relatively higher. By considering that sample k has the highest R-square and using this sample for the training, the result of MAE on each test sample in the whole dataset could be best summarized in Fig. 4.

Fig. 4 shows that, if the system is trained by sample k and be tested on the other data, the MAE is just below 0.1 and the

MAE is calculated using:

$$MAE = \frac{\sum_i |\text{Real value} - \text{Estimated value}|}{\text{Number of data}} \quad (13)$$

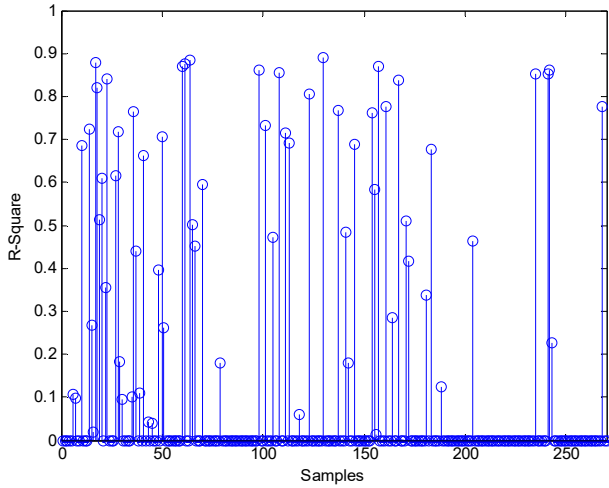


Fig. 3 R square on test data when sample x has been used as the training data

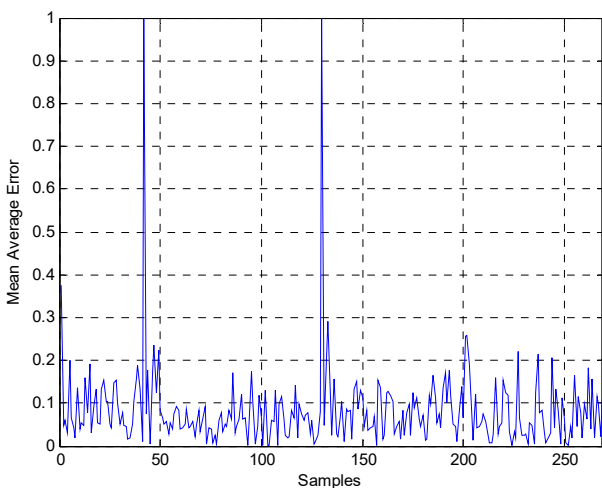


Fig. 4 MAE on test data when sample k has been used as the training data

The estimated parameters, if sample k is used for the training purpose, are presented in (14):

$$\begin{aligned} \gamma(k) &+ a_1 \gamma(k-1) + \dots + a_n \gamma(k-n) \\ &= b_1 u(k-n_m) + \dots + b_n u(k-n-n_m) + \alpha(k) + \\ &+ c_1 \alpha(k-1) + \dots + c_m \alpha(k-m) \end{aligned} \quad (14)$$

$$A(z) = 1 - 1.504z^{-1} - 0.4602z^{-2}$$

$$B(z) = 6.227e^{-0.9}$$

$$C(z) = 1 - 0.9948z^{-1}$$

As seen in (14), the values for the input data are very low and that is due to the higher values of MGT in respect to the gauge values. The real data versus the estimated data are plotted in Fig. 5 showing the accuracy of the model.

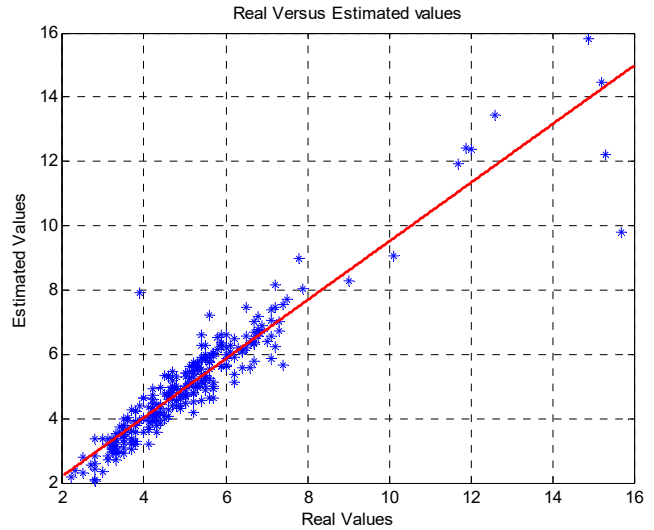


Fig. 5 Real data versus estimated data on test samples

To summarize the findings of this study, the most important factors influencing the tram track degradation are summarized in Table I.

TABLE I
 STATISTICAL PARAMETERS OF MODEL

Parameters	Test data
R square	0.8901
MSE	0.5630
MAE	0.0796

As seen in Fig. 5 and Table I, the R-square indicates reveals that the model explains most of the variability of the response data nearby its mean and is accurate enough in predicting tram track degradation.

IV. CONCLUSION

It is highly necessary for maintenance authorities to acquire a superior level of understanding on how the light rail tracks degrade overtime according to different factors. Also, knowing how the tracks behave in the long run will provide them the advantage of predicting the maintenance work accurately. Having done this, it decreases the amount of money that needs to be spending unnecessarily in maintenance due to human error. In this paper, a dynamic model is put forward to model tram track degrading. At the first part, we have shown that there is a meaningful relationship between gauge and MGT. Then, ARMAX is used due to the complexity of data and their noise. The model is trained by different samples of the data and tested on the other parts of these samples. Although the test data include a high percentage of noise and many fluctuations, the model could follow the gauge value with very low error percentage.

In our model, we do not consider other parameters such as twist and cant, and we only focuses on curves segments. In a further work, those parameters can be analysed on other segments such as straights and crossovers.

ACKNOWLEDGMENT

The Authors would like to thank the support and collaboration of Yarra Trams Railway Infrastructure for providing the data for this research.

REFERENCES

- [1] Y.-J. Zhang, M. H. Murray, and L. Ferreira, "Modelling rail track performance: an integrated approach", *Transport Journal*, 2000, pp.187-194.
- [2] Q. He, H. Li, D. Bhattacharjya, D. P. Parikh, and A. Hampapur, "Track geometry defect rectification based on track deterioration modelling and derailment risk assessment," *Journal of the Operational Research Society*, vol. 66, 2015, pp. 392-404.
- [3] S. Jovanovic, H. Guler, and B. Coko, "Track degradation analysis in the scope of railway infrastructure maintenance management systems," *Gradevinar*, vol. 67, 2015, pp. 247-257.
- [4] A. Iman Soleimanmeigouni, A. Uday Kumar, "Track geometry degradation and maintenance modelling," A review *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 2016.
- [5] Yarra Trams, Facts and figures, Yarra Trams, Melbourne, 2017.
- [6] B. Lichtberger, Track Compendium, Eurailpress Tetzlaff-Hestra GmbH & CO. KG, 2005.
- [7] C. Esveld, "Modern railway track," MRT-productions Zaltbommel, The Netherlands, 2001.
- [8] R. Fröhling, "Wheel/rail interface management in heavy haul railway operations—applying science and technology," *Vehicle System Dynamics*, vol. 45, 2007, pp. 649-677.
- [9] N. Lyngby, P. Hokstad, and J. Vatn, "RAMS management of railway tracks," *Handbook of perfromability engineering*, Springer, 2008.
- [10] L. Wilson, "Optimization of rail/wheel lubrication," Doctoral Thesis, Queensland University of Technology, 2006.
- [11] A. R. Andrade, P. F. Teixeira, "Statistical modelling of railway track geometry degradation using Hierarchical Bayesian models," *Reliability Engineering and System Safety*, vol. 142, 2015, pp. 169-183.
- [12] J. Bernardo, "Bayesian statistics," Updated and abridged version of the Chapter 'Bayesian statistics' published in *Probability and Statistics* (R. Viertl, Ed.) of the Encyclopedia of Life Support Systems (EOLSS), Oxford, UK: UNESCO, 2003.
- [13] I. Ntzoufras, "Bayesian modeling in WinBugs," Hoboken, NJ, USA: John Wiley & Sons, Inc., 2009.
- [14] L. M. Quiroga, E. Schnieder, "Monte Carlo simulation of railway track geometry deterioration and restoration," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 226, 2011, pp. 274-282.
- [15] Y. Shafahi, R. Hakhamaneshi, "Application of a maintenance management model for Iranian railways based on the Markov chain and probabilistic dynamic programming," *International Journal of Science and Technology. Transaction A: Civil Engineering*, vol. 16, 2009, pp. 87-97.
- [16] D. Prescott, J. Andrews, "Modelling maintenance in railway infrastructure management," Reliability and Maintainability Symposium (RAMS), 2013 Proceedings-Annual, 2013, IEEE, pp. 1-6.
- [17] C. Vale, S. M. Lurdes, "Stochastic model for the geometrical rail track degradation process in the Portuguese railway Northern Line," *Reliability Engineering & System Safety*, vol. 116, 2013, pp. 91-98.
- [18] H. Guler, "Prediction of railway track geometry deterioration using artificial neural networks: a case study for Turkish state railways," *Structure and Infrastructure Engineering*, 2014, pp. 614-626.
- [19] J. Sadeghi, H. Askarinejad, "Application of neural networks in evaluation of railway track quality condition," *Mechanical Science Technology* (26), 2012, pp. 113-122.
- [20] Y. Shafahi, P. Masoudi, and R. Hakhamaneshi, "Track Degradation Prediction Models, Using Markov Chain, Artificial Neural and Neuro-Fuzzy Network," 2008.
- [21] S. Moridpour, R. Hesami, "Degradation and performance specification of Melbourne tram tracks," International Conference on Transportation Information and Safety (ICTIS), 25-28 June, 2015.
- [22] S. Moridpour, E. Mazloumi, and R. Hesami, "Application of Artificial Neural Networks in Predicting the Degradation of Tram Tracks Using Maintenance Data", *Applied Big Data Analytics in Operations Management*. IGI Global, U.S, 2016.
- [23] M. Yousefikia, S. Moridpour, S. Setunge, and E. Mazloumi, "Modeling degradation of tracks for maintenance planning on a tram line," Engineering and Technology Publishing, U.S, 2014.
- [24] C. Gao, M. Jiang, J. Y. Huang, and X. F. Wang, "Degradation path modeling method based on time series analysis," *Applied Mechanics and Materials*, 2012.
- [25] M. Khashei, M. Bijari, "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting," *Applied Soft Computing*, vol. 11(2), 2011, pp. 2664-2675.
- [26] A. Samé, F. Chamroukhi, G. Govaert, and P. Aknin, "Model-based clustering and segmentation of time series with changes in regime," *Advances in Data Analysis and Classification*, vol. 5(4), 2011, pp. 301-321.
- [27] L. Wang, H. Zhang, and H. Xue, "Life Prediction based on Degradation Amount Distribution using Composite Time Series Analysis," *IERI Procedia*, 1, 2012, pp. 217-224.
- [28] A. Wijeweera, H. To, M. B. Charles, and K. Sloan, "A time series analysis of passenger rail demand in major Australian cities," *Economic Analysis and Policy*, vol. 44(3), 2014, pp. 301-309.
- [29] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, 2003, pp. 159-175.
- [30] O. Nelles, "Nonlinear system identification: from classical approaches to neural networks and fuzzy models," Springer Science & Business Media, 2013.