The Seventh Asian-Pacific Symposium on Structural Reliability and Its Applications (APSSRA2020)
October 4–7 2020, Tokyo, Japan
T. Takada, I. Yoshida & T. Itoi (editors)

Efficient Data-Driven Prediction Approach for Track Break Zones Based on Extreme Learning Machine

Shui-Hua Jiang[1], Hao-Long Zhang[2], Jinsong Huang[3], and Jinhui Li[4]

[1]*School of Civil Engineering and Architecture, Nanchang University, 999 Xuefu Road, Nanchang 330031, P. R. China. Email: sjiangaa@ncu.edu.cn*
[2]*School of Civil Engineering and Architecture, Nanchang University, 999 Xuefu Road, Nanchang 330031, P. R. China. Email: B7644@ioffice.email*
[3]*Discipline of Civil, Surveying and Environmental Engineering, Faculty of Engineering and Built Environment, The University of Newcastle, Callaghan, NSW 2308, Australia. Email: jinsong.huang@newcastle.edu.au*
[4]*Department of Civil and Environmental Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, P. R. China. Email: jinhui.li@hit.edu.cn*

Abstract:
Accurately predicting the rail track degradation and breaks is the key for formulating the maintenance and renewal strategies of the railway tracks. However, few attempts have been made on the prediction of track break time and zones. This paper proposes a data-driven prediction approach of track break zones based on extreme learning machine. Real-time monitoring data of vertical deformation along the railway tracks are treated as the input, while the break degree that is closely related to the vertical deformation is defined and taken as the output of the system. Based on this, a prediction model of track break zones is constructed using a multi-objective optimization-based extreme learning machine. The difficulties in the determining an optimal number of hidden layer nodes and a rational activation function are overcome by using an orthogonal test method. The real rail break and non-break events of the Instrumented Coal Wagons (ICW) lines in the Hunter Valley of New South Wales, Australia are investigated to illustrate the effectiveness of the proposed approach. The results indicate that the proposed approach can well predict the break zones along the railway tracks based on a large-scale monitoring data. It is expected to be extended to actual projects for predicting the track breaks in the coming days or even several weeks based on the latest monitoring data.

Keywords: railway track, data-driven, degradation prediction, track break, extreme learning machine.

1. Introduction

High speed train transport infrastructure which accommodates the movement of people and the mass transit constitutes the backbone of every nation's economic and social development. However, heavy tonnages, small gaps and complex actions of forces on the tracks make the high speed railway tracks be prone to suffer from the wear, weld or insulated damages (e.g., Xu et al., 2013; Stenström et al., 2016). As a result, not only the track lifespan will be reduced and the railway inspection, maintenance and renewal costs will be increased, but also traffic accidents can be happen once the railway tracks are seriously damaged or broken. Therefore, accurately predicting the development of the track irregularities and the break zones along the railway tracks is the key for formulating the maintenance and renewal strategies (e.g., grinding, renovation and replacement) and reducing the occurrence of traffic accidents.

With the development of the information technology and artificial intelligence, it is feasible to predict the track break time and zones via performing real-time high-precision monitoring and inspection (e.g., Lidén, 2015; Khajehei et al., 2019; Falamarzi et al., 2019a,b). In recent years, significant advances have been made in the prediction of the rail track degradation and breaks. For example, Koc (2012) designed the geometry of a rail track based on the continuous orbit satellite measurement data using the antennas installed on the moving vehicle. Vale and Lurdes (2013) presented a stochastic model for characterizing the railway track geometry deterioration process in the Portuguese railway Northern Line. Shen et al. (2015) constructed a detection model of rail head wear using the collected images from a charge-coupled device camera. Jamshidi et al. (2016) developed a probabilistic defect-based risk assessment approach for rail failures in the railway infrastructure. Chudzikiewicz et al. (2017) evaluated the track condition using axle-boxes and car-bodies motions described by acceleration signals on wheelset axle-boxes. Although some methods have been developed to predict the rail track degradation, how to accurately predict the track potential break zones remains an open question and has not been substantially investigated.

As reported in the literature (e.g., Hall et al., 2011; Tan et al., 2017; Martey et al., 2017; Falamarzi et al., 2019a,b), the machine learning approach can provide an effective means for the prediction of rail track degradation. For example, Sadeghi and Askarinejad (2012) established the relationship between the track geometry conditions and the automatic inspection data by the application of neural networks. Lasisi and Attoh-Okine (2018) developed a support vector machine (SVM) model for predicting the threshold of track quality index (TQI) based on the real-time monitoring data. Falamarzi et al. (2019b) adopted the random forest regression, support vector machine and artificial neural network models for the prediction of tram track degradation based on the acceleration data. Lasisi and Attoh-Okine (2020) adopted an unsupervised machine learning method to evaluate the track performance based on the TQI and safety indicators. Typically, the collected inspection data along the railway tracks change spatiotemporally. However, the spatiotemporally varying monitoring data are rarely used for the construction of the prediction models for the rail track degradation or breaks.

This paper aims to propose an extreme learning machine-based approach for predicting the potential break zones along a rail track. The monitoring data of the vertical displacements along the rail track for the Instrumented Coal Wagons (ICW) lines in the Hunter Valley zones of New South Wales, Australia are taken as the input. An index termed track break degree which is closely related with the track deformation is defined and taken as the output of the system. With the proposed approach, the potential break zones along the rail track in the coming days or even several weeks can be predicted based on the latest monitoring data of the track vertical deformation.

2. Approach for predicting track break zones

2.1. Extreme learning machine

Traditional neural networks such as BP neural network and SVM were widely used for predicting the rail track degradation, but they have some limitations, including bad generalization ability, slow convergence, multiple regulating parameters and easily falling into local minimum (e.g., Falamarzi et al., 2019a). In this study, an extreme learning machine is adopted to construct the model for predicting the break zones along a rail track. The extreme learning machine is a feedforward neural network with a single hidden layer (Huang et al., 2004).

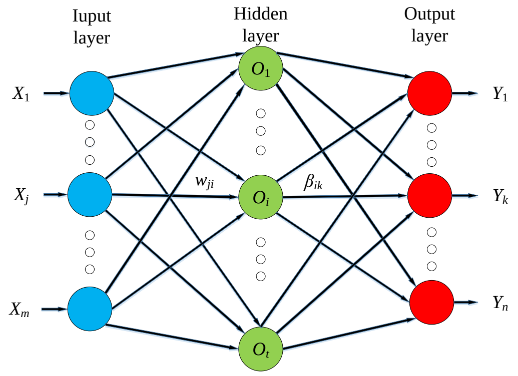


Figure 1. Structure of the extreme learning machine-based feedforward neural network.

The structure of an extreme learning machine-based feedforward neural network with a single hidden layer is shown in Figure 1. It comprises of an input layer, a hidden layer and an output layer. For a given set of learning samples ($\boldsymbol{x}_j$, $\boldsymbol{y}_j$), $\boldsymbol{x}_j = (x_{j1}, x_{j2}, ..., x_{jm})^{\mathrm{T}} \in R^m$, $\boldsymbol{y}_j = (\boldsymbol{y}_{j1}, \boldsymbol{y}_{j2}, ..., \boldsymbol{y}_{jn})^{\mathrm{T}} \in R^n$, $j = 1, 2, \ldots, N$, in which m, n and N are the numbers of the input layer nodes, the output layer nodes and the learning samples, respectively. The feedforward neural network with t hidden layer nodes can be defined as

$$\sum_{i=1}^{t} \boldsymbol{\beta}_i g\left(\boldsymbol{w}_i{}^{\mathrm{T}} \boldsymbol{x}_j + b_i\right) = \boldsymbol{Y}_j \tag{1}$$

where $g(.)$ is the activation function; b_i is the threshold (bias) of the i-th hidden layer neuron; $\boldsymbol{w}_i = \left(w_{1i}, w_{2i}, ..., w_{mi}\right)^{\mathrm{T}}$ and $\boldsymbol{\beta}_i = \left(\beta_{i1}, \beta_{i2}, ..., \beta_{in}\right)^{\mathrm{T}}$ are the weight vectors, in which w_{ji} is the connection weight between the j-th neuron in the input layer and the i-th neuron in the hidden layer, β_{ik} is the connection weight between the i-th neuron in the hidden layer and the k-th in the output layer. The training goal of the extreme learning machine model is to minimize the output error, which can be expressed as

$$\sum_{j=1}^{N} \left\| \boldsymbol{Y}_j - \boldsymbol{y}_j \right\| = 0 \tag{2}$$

Eq. (2) can be rewritten as the following equation for the given $\boldsymbol{w}_i$, $\boldsymbol{x}_j$, b_i:

$$\sum_{i=1}^{t} \boldsymbol{\beta}_i g\left(\boldsymbol{w}_i{}^{\mathrm{T}} \boldsymbol{x}_j + b_i\right) = \boldsymbol{y}_j,\ j = 1, 2, \cdots, N \tag{3}$$

For brevity, Eq. (3) can be simplified as

$$\boldsymbol{H\beta} = \boldsymbol{T} \tag{4}$$

where $\boldsymbol{H}$ is the hidden layer output matrix with dimension of $N \times t$, which represents the output of the hidden layer neurons relative to the input vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, ..., $\boldsymbol{x}_N$, which can be expressed as

$$\boldsymbol{H} = \begin{bmatrix} g(\boldsymbol{w}_1 \boldsymbol{x}_1 + b_1) & \cdots & g(\boldsymbol{w}_t \boldsymbol{x}_1 + b_t) \\ \vdots & \cdots & \vdots \\ g(\boldsymbol{w}_1 \boldsymbol{x}_N + b_1) & \cdots & g(\boldsymbol{w}_t \boldsymbol{x}_N + b_t) \end{bmatrix}_{N \times t} \tag{5}$$

$\boldsymbol{\beta}$ and $\boldsymbol{T}$ are given by

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_t \end{bmatrix}_{t \times m} \quad \text{and} \quad \boldsymbol{T} = \begin{bmatrix} \boldsymbol{y}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{y}_N^{\mathrm{T}} \end{bmatrix}_{N \times m} \tag{6}$$

The training of the extreme learning machine model is equivalent to estimating the least square solution of Eq. (4) (e.g., Huang et al., 2006):

$$\min_{\boldsymbol{\beta}} \left\| \boldsymbol{H\beta} - T \right\| \tag{7}$$

The least square solution $\hat{\boldsymbol{\beta}}$ of the above linear system can be estimated as

$$\hat{\boldsymbol{\beta}} = \boldsymbol{H}^{+} \boldsymbol{T} \tag{8}$$

where $\boldsymbol{H}^{+}$ is the Moore-Penrose generalized inverse of $\boldsymbol{H}$. The norm of $\hat{\boldsymbol{\beta}}$ that is obtained using Eq. (8) will be the smallest and unique (Huang et al., 2006).

Determining an optimal number t of hidden layer nodes and a rational activation function is one key step for the construction of extreme learning machine-based prediction model. Three activation functions, namely Sigmoid, Sine and Hardlim, are frequently used in the extreme learning machine (e.g., Li et al., 2017). The generalization ability of the extreme learning machine could be different even if the same activation function is used. In this study, an orthogonal test method is adopted to determine t and the activation function (e.g., Zhang, 2018). The design space of t is discretized into 20 possible values ranging from 10 to 200 with an increment of 10. Achieving a minimal root mean square error (RMSE) is selected as an evaluation standard to find the optimal combination of t and the activation function. The RMSE is defined as (e.g., Su et al., 2020)

$$\mathrm{RMSE} = \sqrt{\frac{1}{n_r}\sum_{t=1}^{n_r}\left(D_{actual} - D_{predicted}\right)^2} \tag{9}$$

where n_r is the number of input variables; D_{actual} is the actual value of output response; $D_{predicted}$ is the predicted value of output response.

2.2. Definition of break degree

The ICW train lines for the coal transportation in the Hunter Valley zones of New South Wales, Australia with a large-scale monitoring data are taken for illustration. The extreme learning machine is used to construct the prediction model of the break zones along a rail track. The monitoring data of the vertical displacements of the rail track, namely LP_1~LP_4, are taken as the model input. Figure 2 presents a schematic diagram of 4 pairs of ICW train wheels and the layout of monitoring points.

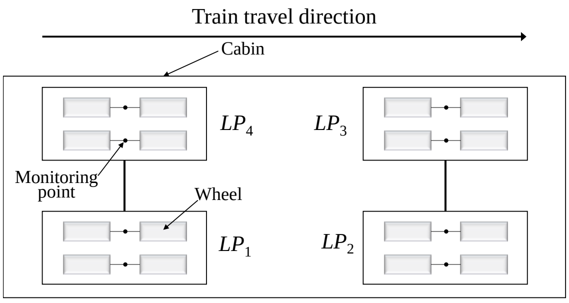
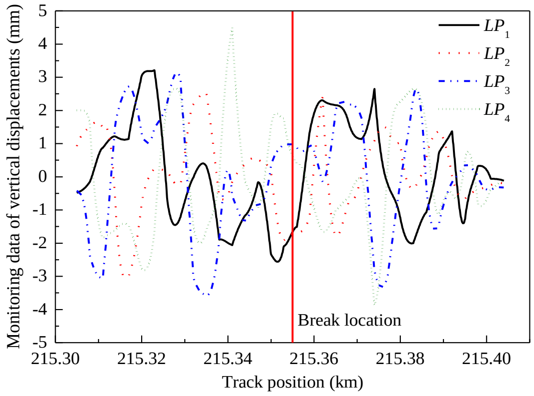


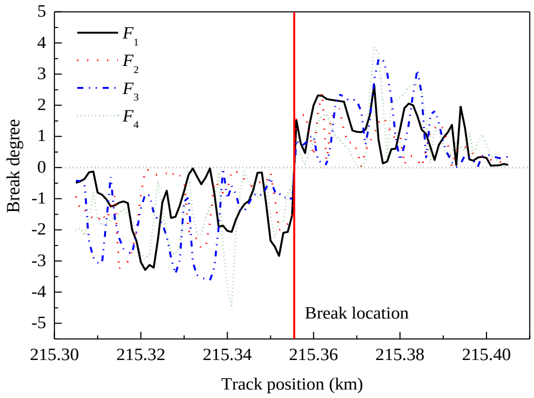
Figure 2. Layout of the monitoring points of vertical displacements for ICW train.

An index termed break degree that is closely related to the vertical deformation of the rail track will be treated as the model output. The break degree is defined as follows: (1) The break degrees corresponding to the railway track before the potential break locations are set as negative values, and their magnitudes are equal to the absolutes of the corresponding vertical displacements. (2) The break degrees corresponding to the railway track after the potential break locations are set as positive values, and their magnitudes are also equal to the absolutes of the corresponding vertical displacements. (3) The break degrees corresponding to the potential break locations are set as zero.

Take the rail track break event of 80798 for an example, Figure 3(a) shows the variations of the vertical displacements LP_1~LP_4 measured by the ARTC03 train on April 13, 2017. The horizontal ordinate represents the track position (km) while the vertical rdinate represents the vertical displacements of the rail track (mm). The study interval is 215.305~215.405 km. The track was broken at the middle of the study interval on April 26, 2017. Figure 3(b) shows the variations of the corresponding break degrees F_1~F_4 with the track position. As defined, the values of the break degrees associated with the non-break zones are negative, and those for the zones after the break locations are positive. The magnitudes are exactly equal to the absolutes of the corresponding vertical displacements.



(a) Monitoring data of vertical displacements LP_1~LP_4



(b) Break degrees F_1~F_4

Figure 3. Variations of the monitoring data of vertical displacements and the corresponding break degrees with the track position.

3. Model construction and verification

3.1. Construction of prediction model

The monitoring data of vertical displacements from the 30 rail track break events and 20 non-break events are selected as the training samples. As mentioned earlier, the number (t) of the hidden layer nodes and the activation function need to be determined prior to the construction of the prediction model. Figure 4 shows the RMSE values for different combinations of t and activation functions calculated from the orthogonal test method. It can be observed that the lowest value of RMSE = 0.253 corresponds to the combination of t =110 and activation function of “Sin”. Therefore, t=110 and “Sin” function are adopted in the following model construction and verification. Note that the number (m) of the input layer nodes is set to equal that (n) of the output layer nodes in this study.

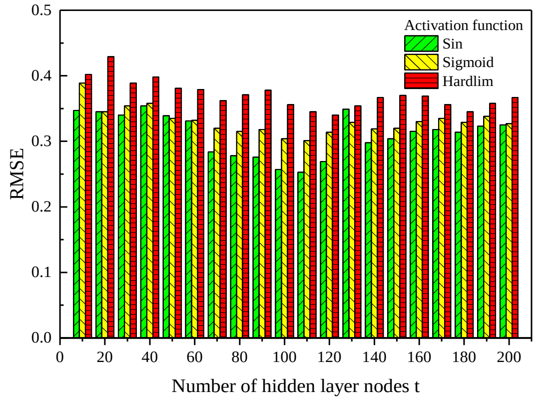
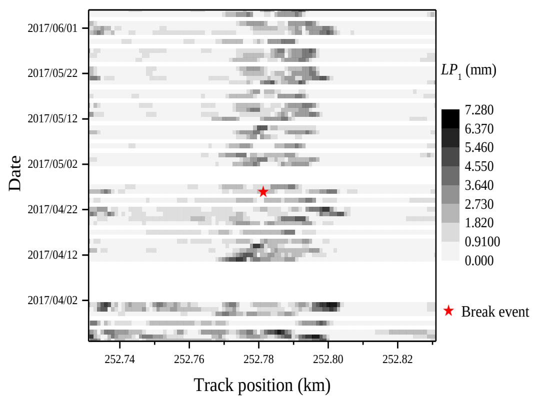


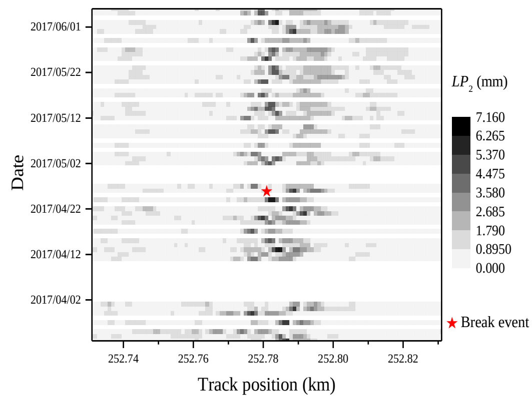
Figure 4. Orthogonal test results for different combinations of the number of hidden layer nodes and the activation function.

3.2. Model testing for break events

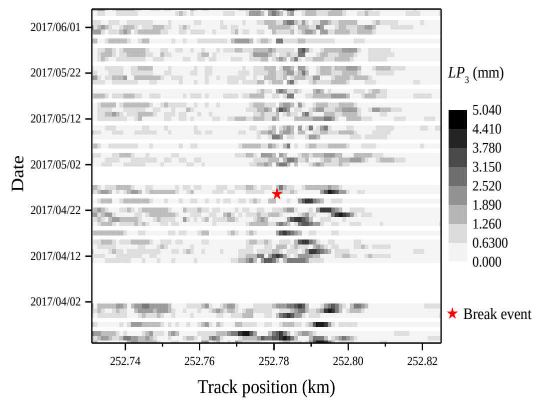
The monitoring data of vertical displacements from the other 10 rail track break events and 10 non-break events are employed for the model testing. For example, the break event of 80799 is taken to test the prediction model. The rail track broke at 252.781 km on April 26, 2017 for this break event. The study interval is 252.731~252.831 km and the period is from March 24, 2017 to June 5, 2017. Figure 5 presents variations of the vertical displacements monitored by the ARTC03 train with the time and track position. As observed from Figure 5, an asymptotic trend toward the track break in the neighbor of the track position of 252.781 km can be observed from March 24, 2017 to April 25, 2017. The vertical displacements particularly LP_1 and LP_4 in these regions change significantly. The potential break zones along the rail track can be roughly inferred according to the variation trends of the vertical displacements. Based on the analysis of the spatiotemporally varying data in Figure 5, the maintenance time and strategies for preventing the track break can be determined. The model prediction accuracy of the break zones along the rail track can be further increased if more updated monitoring data are incorporated for constructing the prediction model.



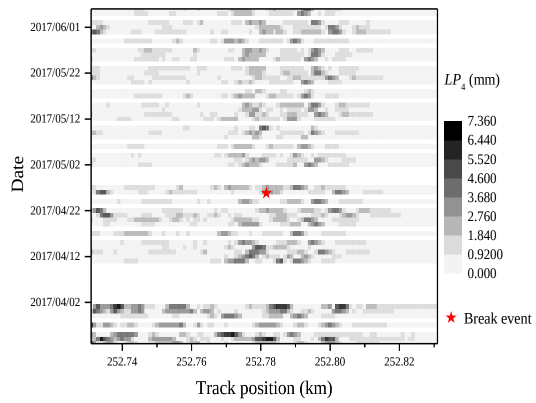
(a) LP_1



(b) LP_2



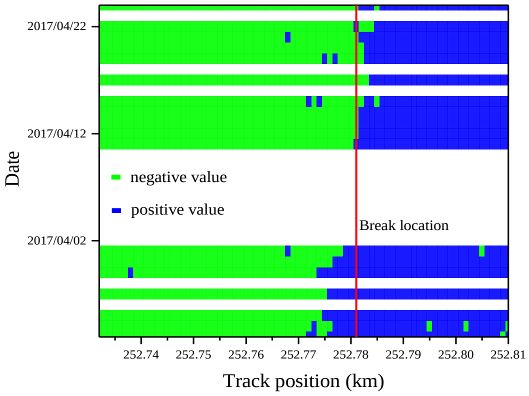
(c) LP_3



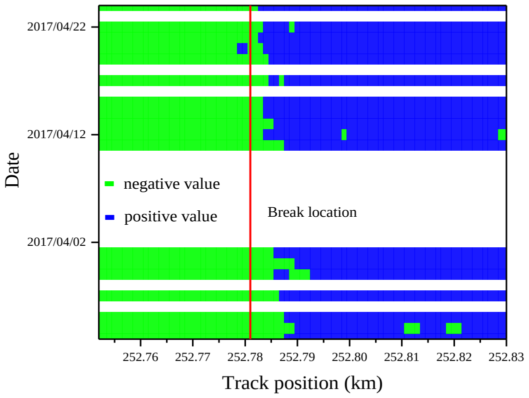
(d) LP_4

Figure 5. Variations of the monitoring data of vertical displacements with time and track position for the break event of 80799.

Two sets of the monitoring data of vertical displacements monitored from March 24, 2017 to April 25, 2017 acquired at different track positions are taken as the testing samples to validate the prediction model. Figures 6(a) and (b) present the corresponding prediction results of the break degrees. To enable a better distinction of the results, the regions with the break degrees being negative values are assigned as green, while the regions with the break degrees being positive values are assigned as blue. The red line represents the real track break position. It can be seen from Figure 8 that the proposed approach can well predict the potential break zones along railway tracks even if different model inputs are used. It confirms the effectiveness of the proposed approach.



(a) Model testing results for the first set of testing samples



(b) Model testing results for the second set of testing samples

Figure 6. Testing results of the prediction model for the break event of 80799.

Moreover, the computational time taken by the proposed approach for the model training and testing is about 60 and 2.2 seconds, respectively, on a desktop with 8 GB RAM and one Intel Core E5 CPU clocked at 3.2 GHz. It indicates that the proposed approach is quite efficient to predict the potential break zones along the rail track. Due to the complexity of the model input and the high dimensionality (79) of learning samples, the under-fitting may be induced for the extreme learning machine model because of complicated least squares regression (e.g., Bartlett, 1998; Chapelle, 2002). These limitations shall be overcome in the future study.

3.3. Model testing for non-break events

In this section, the non-break event of 342 is further selected to test the prediction model. Figure 7 shows the vertical displacements of the rail track monitored on May 25, 2017, which are used as the testing samples. Figure 8 presents the prediction result of the break degrees. The average $F_{average}$ of the four break degrees is also plotted in Figure 8 to judge where the track break zones are. It can be seen from Figure 8 that the obtained four break degrees are almost smaller than 0 along the rail track, and the average $F_{average}$ of these four break degrees are also negative values. It indicates the predicted results are consistent with the engineering practice wherein this rail track was not been damaged. This further illustrates the effectiveness of the prediction approach.

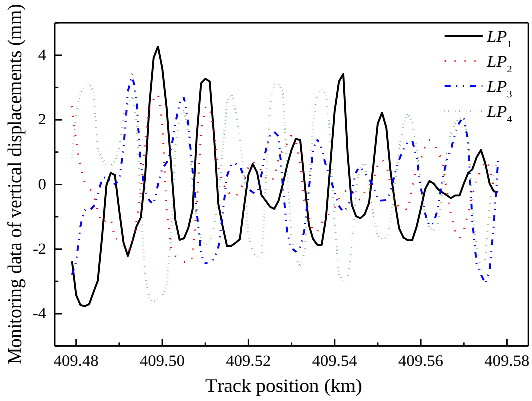


Figure 7. Variation of the monitoring data of vertical displacements with the track position for the non-break event of 342.

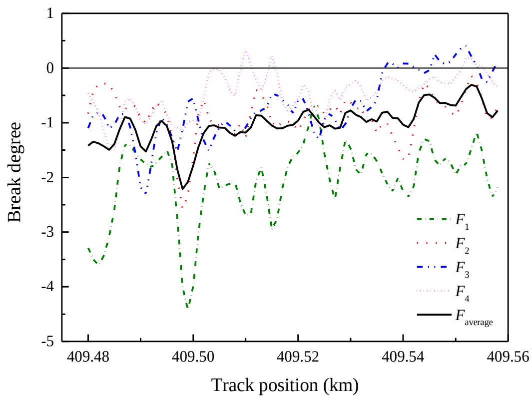


Figure 8. Testing results of the prediction model for the non-break event of 342.

4. Conclusions

In this paper, an extreme learning machine-based approach for predicting the break zones along a rail track is proposed. The monitoring data of vertical displacements of the rail track are taken as the input. The break degree that is closely related to the vertical displacements is defined and taken as the output of the system. The ICW lines with a large-scale monitoring data in the Hunter Valley zones of New South Wales, Australia are used for illustrate the effectiveness of the proposed approach. The monitoring data of vertical displacements from the real 40 track break events and 30 non-break events are used as training samples and testing samples to construct and test the prediction model. It is confirmed that the proposed approach can efficiently predict the potential break zones along the rail track. It can be extended to actual projects for predicting the track breaks in the coming days or even several weeks using the latest monitoring data. Based on these, the maintenance and renewal strategies for preventing the track degradation and breaks can be formulated. Additionally, the variation trends of the monitored data of vertical displacements of the rail track can be used to roughly determine the track irregularities, but they cannot be directly used to predict the potential break zones along the railway track.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Project No. 41867036, 41972280) and the Natural Science Foundation of Jiangxi Province (Project No. 2018ACB21017, 20181ACB20008, 20192BBG70078). The financial supports are gratefully acknowledged.

References

Bartlett, P.L. 1998. The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network. *IEEE Transactions on Information Theory*, 44(2): 525–536.

Chapelle, O., Vapnik, V. and Bengio, Y. 2002. Model Selection for small sample regression. *Machine Learning*, 48(1): 9–23.

Chudzikiewicz, A., Bogacz, R., Kostrzewski, M. and Konowrocki, R. 2017. Condition monitoring of railway track systems by using acceleration signals on wheelset axle-boxes. *Transport*, 33(2): 555–566.

Falamarzi, A., Moridpour, S., and Nazem, M. 2019a. A review of rail track degradation prediction models. *Australian Journal of Civil Engineering*, 17(2): 152–166.

Falamarzi, A., Moridpour, S., and Nazem, M. 2019b. Development of a tram track degradation prediction model based on the acceleration data. *Structure and Infrastructure Engineering*, 15(10): 1308–1318.

Huang, G.B., Zhu, Q.Y. and Siew, C.K. 2004. Extreme learning machine: a new learning scheme of feedforward neural networks. *Neural networks*, 2: 985–990.

Huang, G.B., Zhu, Q.Y. and Siew, C.K. 2006. Extreme learning machine: *Theory and applications. Neurocomputing*, 70(1-3): 489–501.

Hall, T., Beecham, S., Bowes, D., Gray, D. and Counsell, S. 2011. A systematic literature review on fault prediction performance in software engineering. *IEEE Transactions on Software Engineering*, 38(6): 1276–1304.

Jamshidi, A., Roohi, S.F., Núñez, A., Babuska, R., Schutter, B.D., Dollevoet, R. and Li, Z. 2016. Probabilistic defect-based risk assessment approach for rail failures in railway infrastructure. *International Federation of Automatic Control*, 49(3): 73–77.

Khajehei, H., Ahmadi, A., Soleimanmeigouni, I. and Nissen, A. 2019. Allocation of effective maintenance limit for railway track geometry. *Structure and Infrastructure Engineering*, 15(12): 1597–1612.

Koc, W. 2012. Design of rail-track geometric systems by satellite measurement. *Journal of Transportation Engineering*, 138(1): 114–122.

Lidén, T. 2015. Railway infrastructure maintenance - a survey of planning problems and conducted research. *Transportation Research Procedia*, 10(7): 574–583.

Lasisi, A. and Attoh-Okine, N. 2018. Principal components analysis and track quality index: A machine learning approach. *Transportation Research Part C: Emerging Technologies*, 91: 230–248.

Lasisi, A. and Attoh-Okine, N. 2020. An unsupervised learning framework for track quality index and safety. *Transportation Infrastructure Geotechnology*, 7(1): 1–12.

Li, H., Yu, H., Tian, Z. and Li, B. 2017. Prediction of deterioration of rolling bearings based on binary multi-scale entropy. *China Mechanical Engineering*, 28(20): 2420–2425.

Martey, E.N., Ahmed, L. and Attoh-Okine, N. 2017. Track geometry big data analysis: A machine learning approach. In *2017 IEEE International Conference on Big Data (Big Data). IEEE*, pp. 3800–3809.

Sadeghi, J. and Askarinejad, H. 2012. Application of neural networks in evaluation of railway track quality condition. *Journal of Mechanical Science and Technology*, 26 (1): 113–122.

Shen, J.J., Liu, Z.J., Wu, T. and Lv, M. 2015. A rail abrasion detection system based on image processing. *Applied Mechanics & Materials*, 744-746: 1213–1216.

Stenström, C., Norrbin, P., Parida, A. and Kumar, U. 2016. Preventive and corrective maintenance–cost comparison and cost–benefit analysis. *Structure and Infrastructure Engineering*, 12(5): 603–617.

Su, X., An, J., Zhang, Y. and Zhu, P. 2020. Prediction of ozone hourly concentrations by support vector machine and kernel extreme learning machine using wavelet transformation and partial least squares methods. *Atmospheric Pollution Research*, 11(6): 51–60.

Tan, C.W. Webb, G.I., Petitjean, F. and Reichi, P. 2017. Tamping effectiveness prediction using supervised machine learningtechniques. *First International Conference on Rail Transportation*, pp. 1010–1023.

Vale, C. and Lurdes, M.S. 2013. Stochastic model for the geometrical rail track degradation process in the portuguese railway northern line. *Reliability Engineering and System Safety*, 116: 91–98.

Xu, G., Shi, T., Ren, S., Han, Q. and Wang, D. 2013. Development of the on-board track inspection system based on computer vision. *China Railway Science*, 34(1): 139–144.

Zhang, Z. 2018. Deformation prediction of bank slope based on chaos optimization extreme learning machine. *Hydroelectric Power*, 44 (12): 39–42.